



金融计量学：原理与应用

作者：徐秋华、罗荣华

单位：西南财经大学金融学院

时间：January 17, 2023

版本：0.3

联系方式：xuqh@swufe.edu.cn



All men by nature desire to know. -- Aristotle

前言

作者

于成都·西财柳林校区

2022 年 4 月

目录

前言	i
第一部分 收益率及其性质	1
1 收益率及其分布特征	2
1.1 收益率的定义与计算	2
1.2 资产收益率的典型特征	8
1.3 收益率的分布	11
1.4 有效市场假说和收益率的可预测性*	16
1.5 总结	27
附录 1: 对数正态分布的期望和方差	28
附录 2: 正态性检验	28
第一章 习题	34
第二部分 线性时间序列模型	35
2 一元线性时间序列模型	36
2.1 自相关函数	36
2.2 移动平均模型	36
2.3 自回归模型	36
2.4 自回归移动平均模型	36
3 多元线性时间序列模型	37
4 单位根与协整	38
5 状态空间模型及卡尔曼滤波*	39
6 混频数据抽样回归	40
第三部分 非线性模型	41
7	42
8	43
9 深度学习模型	44
第四部分 波动率模型	45

10 一元波动率模型	46
11 多元波动率模型	47
12 随机波动率模型 *	48
第五部分 量化风险管理	49
13 单资产风险测度	50
14 极值理论 *	51
15 Copula	52
16 系统风险测度 *	53
参考文献	54
附录	59
A EM 算法 *	60
B 贝叶斯估计 *	61

第一部分 收益率及其性质

第一章 收益率及其分布特征

内容提要

- 简单收益率 1.1
- 对数收益率 1.1
- 收益率的典型特征 1.2
- 收益率的分布 1.3
- 有效市场假说 1.4
- 收益率的可预测性 1.4
- 鞅 1.4
- 随机游走 1.4
- Cowles-Jones 比率 1.1
- 方差比检验 1.4

收益率是金融学的核心概念之一，是资产持有者因承担风险而应得的回报。大多数金融时间序列分析都以资产的收益率作为研究对象，这是因为收益率具有良好的统计性质，例如：收益率通常是平稳的，而价格是非平稳的。此外，收益率是投资获利机会的直接体现，而且收益率与量纲无关。本章第一节将介绍简单收益率和对数收益率；第二节讨论资产收益率数据的典型特征；第三节讨论收益率的分布；本章最后将讨论收益率的可预测性和有效市场假说。想深入了解本章内容可以进一步参考 [Campbell et al. \(1997\)](#); [Linton \(2019\)](#); [Fan et al. \(2015\)](#); [Jondeau et al. \(2007\)](#); [Tsay \(2014\)](#)。

1.1 收益率的定义与计算

令 P_t 表示资产在 t 时刻的价格，我们接下来给出各种资产收益率的定义，并讨论这些收益率的主要性质：

简单收益率

首先假设资产不支付红利，则持有资产一期的简单收益率定义为：

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1.$$

R_t 表示将资产从第 $t-1$ 期持有到第 t 期所获得的利润率。通常，我们将收益率记为 $R_t = 100R_t\%$ ，例如，若 $R_t = 0.0175$ ，我们记为 $R_t = 1.75\%$ 。持有资产一期的总收益 (gross return) 定义为：

$$\frac{P_t}{P_{t-1}} = R_t + 1.$$

类似地，我们可以定义多期简单收益率。若持有资产 k (≥ 1) 期， k 期简单收益率定义为：

$$R_t[k] = \frac{P_t - P_{t-k}}{P_{t-k}} = \frac{P_t}{P_{t-k}} - 1.$$

相应的 k 期总收益定义为：

$$\frac{P_t}{P_{t-k}} = 1 + R_t[k].$$

容易验证：

$$\begin{aligned}
 1 + R_t[k] &= \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}} \\
 &= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}) \\
 &= \prod_{j=0}^{k-1} (1 + R_{t-j}),
 \end{aligned}$$

即，持有资产 k 期的总收益等于持有期内的 k 个单期总收益的乘积。当单期简单收益率 R_{t-k+1}, \dots, R_t 很小时，我们有：

$$R_t[k] \approx R_t + R_{t-1} + \cdots + R_{t-k+1}.$$

根据上述定义，资产的简单收益率与持有资产的期数有关。在比较不同资产的收益率时，我们通常使用的是年化收益率，其定义为：

$$\text{Annualized}\{R_t[k]\} = \left[\prod_{j=0}^{k-1} (1 + R_{t-j}) \right]^{1/k} - 1 = \exp \left[\frac{1}{k} \sum_{j=0}^{k-1} \ln(1 + R_{t-j}) \right] - 1,$$

其中， k 为持有资产的年数。注意到：若 $x \rightarrow 0$ ， $\ln(1+x) \approx x$ ， $\exp(x) \approx 1+x$ 。因此，当单期简单收益率很小时，

$$\text{Annualized}\{R_t[k]\} \approx \frac{1}{k} \sum_{j=0}^{k-1} R_{t-j}.$$

关于资产组合的简单收益率，我们有如下结论：

定理 1.1. 资产组合简单收益率的可加性

令 w_{1t}, \dots, w_{nt} 分别表示资产组合中的 n 个资产在 t 时刻的组合权重，则资产组合的简单收益率 $R_{p,t}$ 为：

$$R_{p,t} = w_{1t}R_{1t} + w_{2t}R_{2t} + \cdots + w_{nt}R_{nt}.$$



证明 令 m_i 表示资产组合中第 i 个资产的数量，则 t 时刻资产的价格为 $V_t = m_1P_{1t} + m_2P_{2t} + \cdots + m_nP_{nt}$ ，第 i 个资产 t 时刻所占比重为 $w_{it} = m_iP_{it}/V_t$ 。若将此资产组合持有到第 $t+1$ 期，则其在 $t+1$ 时刻的价格为：

$$\begin{aligned}
 V_{t+1} &= m_1P_{1,t+1} + m_2P_{2,t+1} + \cdots + m_nP_{n,t+1} \\
 &= \frac{w_{1t}V_t}{P_{1t}}P_{1,t+1} + \cdots + \frac{w_{nt}V_t}{P_{nt}}P_{n,t+1} \\
 &= \frac{w_{1t}V_t}{P_{1t}}P_{1t}(1 + R_{1,t+1}) + \cdots + \frac{w_{nt}V_t}{P_{nt}}P_{nt}(1 + R_{n,t+1}) \\
 &= w_{1t}V_t(1 + R_{1,t+1}) + \cdots + w_{nt}V_t(1 + R_{n,t+1}) \\
 &= V_t(1 + w_{1t}R_{1,t+1} + \cdots + w_{nt}R_{n,t+1}).
 \end{aligned}$$

于是，

$$R_{p,t+1} = \frac{V_{t+1} - V_t}{V_t} = w_{1t}R_{1,t+1} + \cdots + w_{nt}R_{n,t+1}.$$

□

对数收益率

假设银行的年利率为 r ，年内付息次数是 m 次，年初存入银行的 1 元钱到年底的总价值为 $(1 + \frac{r}{m})^m$ 。若 $m \rightarrow \infty$ (连续复利)，我们有：

$$\lim_{m \rightarrow \infty} \left(1 + \frac{r}{m}\right)^m = e^r.$$

于是， $r = \log(e^r) = \log(e^r) - \log(1)$ ，即， r 为年末价格的对数减去年初价格的对数。一般地，我们将对数收益率定义为：

$$r_t = \log(P_t) - \log(P_{t-1}) = \log(1 + R_t).$$

根据上面的分析，对数收益率也称为连续复利收益率。与简单收益率相比，对数收益率有如下优点：

- 可比较：不同资产的对数收益率可以直接比较，与付息的频率无关；
- 时间可加性：持有资产多期的对数收益率等于持有期内的各个单期对数收益率之和，这是因为

$$\begin{aligned} r_t[k] &= \log(1 + R_t[k]) = \log \left[\prod_{j=0}^{k-1} (1 + R_{t-j}) \right] \\ &= \sum_{j=0}^{k-1} \log(1 + R_{t-j}) = \sum_{j=0}^{k-1} r_{t-j}. \end{aligned}$$

但是，资产组合的对数收益率不具有可加性，具体地，由定理 1.1，

$$R_{p,t} = \sum_{i=1}^n w_{it} R_{it}.$$

对上式两端同时加 1 得到：

$$\begin{aligned} e^{r_{p,t}} &\equiv 1 + R_{p,t} = 1 + \sum_{i=1}^n w_{it} R_{it} \\ &= \sum_{i=1}^n w_{it} + \sum_{i=1}^n w_{it} R_{it} \\ &= \sum_{i=1}^n w_{it} (1 + R_{it}) = \sum_{i=1}^n w_{it} e^{r_{it}}. \end{aligned}$$

对上式两端取对数得到：

$$r_{p,t} = \log \left(\sum_{i=1}^n w_{it} e^{r_{it}} \right) \neq \sum_{i=1}^n w_{it} r_{it}.$$

当简单收益率很小时，我们有：

$$r_t = \log(1 + R_t) \approx R_t.$$

进而，

$$1 + r_{p,t} \approx e^{r_{p,t}} = \sum_{i=1}^n w_{it} e^{r_{it}} \approx \sum_{i=1}^n w_{it} (1 + r_{it}) = 1 + \sum_{i=1}^n w_{it} r_{it},$$

即， $r_{p,t} \approx \sum_{i=1}^n w_{it} r_{it}$ 。同时，由泰勒展开公式，

$$r_t = \log(1 + R_t) = R_t - \frac{R_t^2}{2} + O_p(R_t^3).$$

所以，简单收益率通常会大于对数收益率。

若持有的资产从第 $t-1$ 期到第 t 期之间支付了红利 D_t ，在计算收益率时应该将红利支付纳入考虑。此时，简单收益率和对数收益率的定义分别为：

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1,$$

和

$$r_t = \log(P_t + D_t) - \log P_{t-1}.$$

例 1.1 我们从 WIND 上下载了贵州茅台（600519）从 2018-1-2 至 2022-6-30 的日收盘价数据，利用 R 语言计算贵州茅台在样本时间段内的日、周和月的简单收益率和对数收益率，并对收益率数据进行可视化，具体代码如下：

```
1 mydata <- readxl::read_xlsx("~/GZMT.xlsx")
2 library(xts) # 处理时间序列数据的程序包
3 # 利用xts命令将收盘价转换为xts对象
4 price <- xts(as.numeric(mydata$‘收盘价(元)’),mydata$日期)
5 names(price) <- "price"
6 # 利用quantmod程序包中的命令periodReturn计算收益率，
7 # 类型 ‘arithmetic’ 表示简单收益率
8 drt <- quantmod::periodReturn(price,period = 'daily',type =
  'arithmetic')
9 wrt <- quantmod::periodReturn(price,period = 'weekly',type =
  'arithmetic')
10 mrt <- quantmod::periodReturn(price,period = 'monthly',type =
  'arithmetic')
11 dlogrt <- quantmod::periodReturn(price,period = 'daily',type = 'log')
12 wlogrt <- quantmod::periodReturn(price,period = 'weekly',type = 'log')
13 mlogrt <- quantmod::periodReturn(price,period = 'monthly',type = 'log')
14 plot(price,main = "daily price") # 绘制收盘价的时间序列图
15 plot(dlogrt,main = "daily log return") # 绘制日对数收益率的时间序列图
16 drt <- as.numeric(drt) # 转换为数值型
17 wrt <- as.numeric(wrt)
18 mrt <- as.numeric(mrt)
19 dlogrt <- as.numeric(dlogrt)
20 wlogrt <- as.numeric(wlogrt)
21 mlogrt <- as.numeric(mlogrt)
22 dailydata <- cbind.data.frame(drt,dlogrt) # 合并成数据框
23 weeklydata <- cbind.data.frame(wrt,wlogrt)
24 monthlydata <- cbind.data.frame(mrt,mlogrt)
25 library(tidyverse)
```

```

26 dfigure <- dailydata %>%
27   ggplot(aes(drt, dlogrt)) +
28   geom_point(colour="blue") +
29   labs(x = "简单收益率", y="对数收益率", title = "贵州茅台日收益率") +
30   theme(plot.title = element_text(hjust = 0.5)) + # title居中
31   geom_line(mapping = aes(x=dlogrt, y=dlogrt), colour="red") # 参照线
32
33 wfigure <- weeklydata %>%
34   ggplot(aes(wrt, wlogrt)) +
35   geom_point(colour="blue") +
36   labs(x = "简单收益率", y="对数收益率", title = "贵州茅台周收益率") +
37   theme(plot.title = element_text(hjust = 0.5)) +
38   geom_line(mapping = aes(x=wlogrt, y=wlogrt), colour="red")
39
40 mfigure <- monthlydata %>%
41   ggplot(aes(mrt, mlogrt)) +
42   geom_point(colour="blue") +
43   labs(x = "简单收益率", y="对数收益率", title = "贵州茅台月收益率") +
44   theme(plot.title = element_text(hjust = 0.5)) +
45   geom_line(mapping = aes(x=mlogrt, y=mlogrt), colour="red")
46 # 将三幅图画在同一个画图界面中
47 ggpubr::ggarrange(dfigure, wfigure, mfigure, ncol = 3, nrow = 1)

```

图 1.1 和 1.2 分别绘制了贵州茅台的日收盘价和日对数收益率的时间序列图。从两幅图中可以看出，收益率数据会围绕着某一固定水平上下波动，这一固定水平即为收益率的样本均值，我们称这一现象为“均值回复”（mean reversion），而价格数据是有趋势的非平稳数据。图 1.3 画出了以贵州茅台的简单收益率为横轴，对数收益率为纵轴的散点图，并在其上添加了 $y = x$ 的参照线。从图中我们可以看出，三个频率的简单收益率都不小于相同频率的对数收益率，这与泰勒公式的结论是一致的。

超额收益率

在实证资产定价的研究中，超额收益率是经常用到的收益率概念。超额收益率定义为资产收益率与某种无风险资产收益率之差：

$$Z_t = R_t - R_{0t},$$

$$z_t = r_t - r_{0t},$$

其中， R_{0t} 和 r_{0t} 分别表示无风险资产的简单收益率和无风险资产的对数收益率，因此， Z_t 为资产的简单超额收益率； z_t 为资产的对数超额收益率。这里需要注意的是无风险资

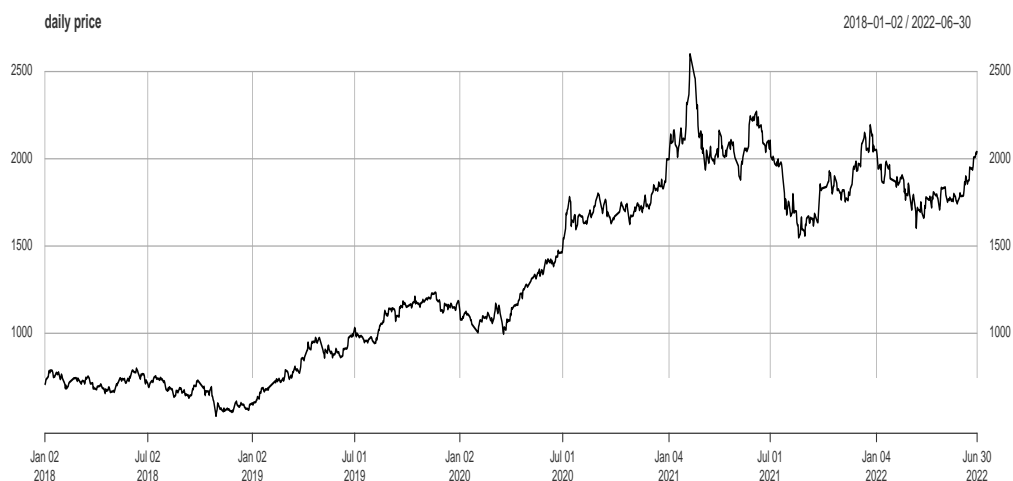


图 1.1: 贵州茅台日收盘价

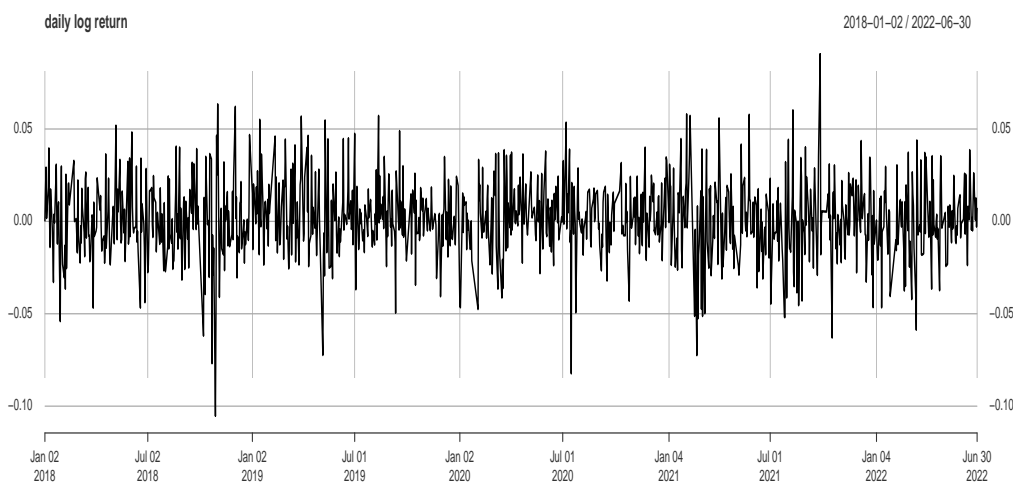


图 1.2: 贵州茅台日对数收益率

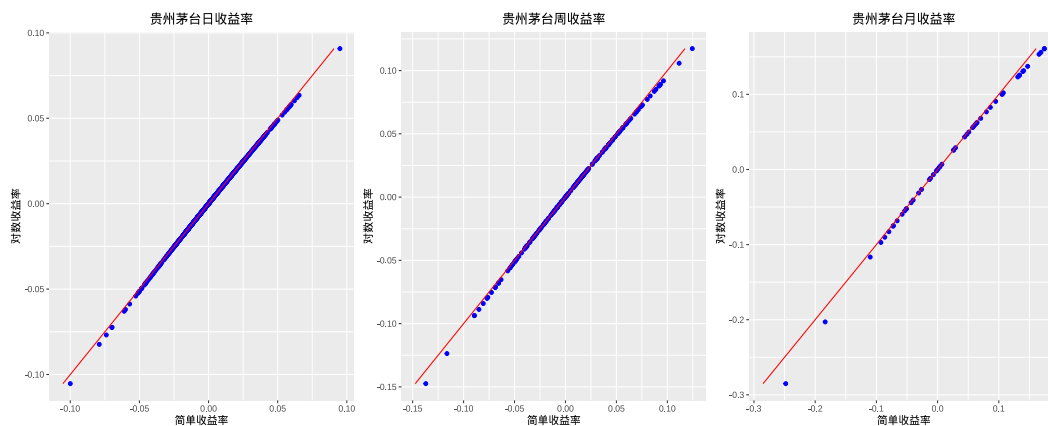


图 1.3: 贵州茅台简单收益率与对数收益率

产并不是完全没有风险的资产，而是指没有违约或信用风险的资产，例如，10 年期国债和 10 年期国开债。

1.2 资产收益率的典型特征

在本章附录 2 中我们指出金融时间序列数据通常具有尖峰、厚尾和非对称等典型的非正态特征。本节进一步讨论金融时间序列数据的一些典型特征。按照惯例，我们分析的金融时间序列数据为对数收益率数据。

对于服从正态分布的随机变量，期望和方差唯一确定了该随机变量的分布。对于不服从正态分布的随机变量，为了解其分布特征，我们往往还会关注它的标准化之后的三阶矩和四阶矩，即，偏度和峰度¹。以例 1.1 中贵州茅台的日对数收益率为例，其偏度和超额峰度分别为 -0.136 和 1.556 ；检验偏度为 0 的 t 检验的 p 值为 0.067 ，在 10% 的水平下拒绝原假设；检验超额峰度为 0 的 t 检验的 p 值为 0，在 1% 的水平下拒绝原假设。Shapiro-Wilk 正态检验的 p 值几乎为 0。图 1.4 绘制了贵州茅台对数收益率的直方图和 Q-Q 图，与第二章中上证指数的结果类似，我们发现贵州茅台的收益率也具有明显的非正态特征。

```

1 # 接例1.1第19行程序
2 library(fBasics)
3 s3 = skewness(dlogrt) # 计算偏度
4 s3
5 [1] -0.135697
6 samplesize = length(dlogrt) # 求样本量
7 t3 = s3/sqrt(6/samplesize) # 检验偏度为0的t统计量，见本章附录2
8 pp3 = 2*(pnorm(t3)) # 计算p值
9 pp3
10 [1] 0.06740325
11 s4 = kurtosis(dlogrt, method = "excess") # 计算超额峰度，即，峰度-3
12 s4
13 [1] 1.555572
14 t4 = s4/sqrt(24/samplesize) # 检验超额峰度为0的t统计量，见本章附录2
15 pp4 = 2*(1-pnorm(t4)) # 计算p值
16 pp4
17 [1] 0
18 shapiro.test(as.numeric(dlogrt))
19 # p-value = 8.639e-10
20
21 mu <- mean(dlogrt)

```

¹ 具体定义见本章附录 2。

```

22 sd <- sd(dlogrt)
23 nrx <- seq(min(dlogrt),max(dlogrt),by=0.001)
24 nry <- dnorm(nrx,mu,sd)
25 normdata <- as.data.frame(cbind(nrx,nry))
26 library(tidyverse)
27 hist1 <- dlogrt %>%
28   ggplot(aes(dlogrt)) +
29   geom_histogram(aes(y = ..density..),fill="yellow",colour="grey30")+
30   labs(x = "贵州茅台对数收益率")+
31   geom_line(mapping = aes(x=nrx,y=nry),data = normdata,colour="blue")
32
33 rtn <- as.numeric(unlist(dlogrt))
34 rtn <- as.data.frame(rtn)
35 qqplot2 <- ggplot(rtn, aes(sample=rtn)) +
36   labs(x = "Normal quantile",y="Quantile of returns")+
37   stat_qq(colour="blue") +
38   stat_qq_line(colour="red")
39
40 ggpubr::ggarrange(hist1, qqplot2, ncol = 2, nrow = 1)

```

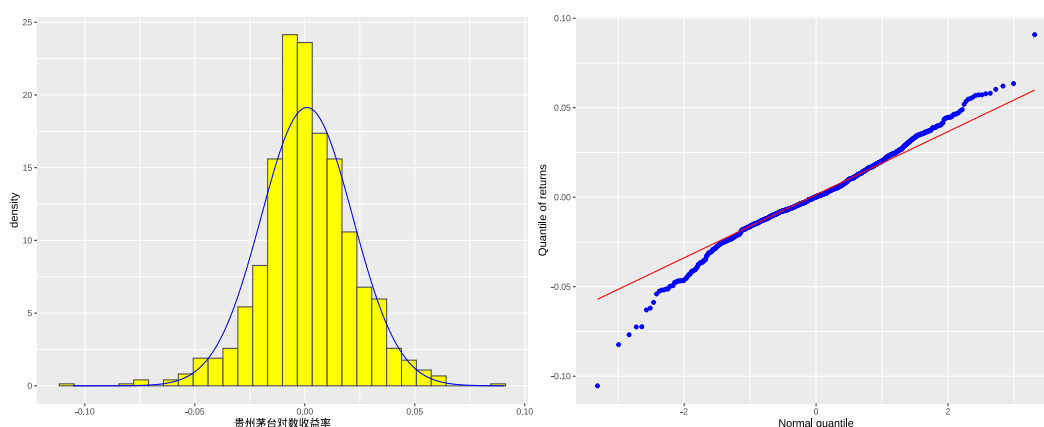


图 1.4: 贵州茅台日对数收益率的直方图（左）和 Q-Q 图（右）

除了尖峰、厚尾和非对称之外，收益率还具有以下主要特征：

- (1) **弱平稳性**：大多数收益率序列可以看作是具有不随时间变化的一阶矩和二阶矩的随机过程的实现（realization）；
- (2) **波动率聚集**：在一定的时间段内，资产价格波动的幅度往往具有持续性，某一期大幅度的价格波动往往伴随着下一期价格波动幅度的放大，对于小幅度的价格波动也有类似现象；

(3) **非线性相依性**：虽然收益率数据往往不存在明显的序列相关性²，但是，经过平方或取绝对值等非线性变换之后的收益率数据呈现出虽小但是显著的序列相关性。

如图 1.5 所示，上证指数收益率围绕着样本均值上下波动，我们可以认为其一阶矩和二阶矩不随时间变化。同时，上证指数收益率也存在波动率聚集的现象。例如，图中左侧红色椭圆标出了高波动率聚集的时间段，右侧红色椭圆标出了低波动率聚集的时间段。这里似乎存在自相矛盾，弱平稳要求收益率的二阶矩不随时间变化，二阶矩中的方差在金融中被用来衡量波动的幅度。既然方差不随时间变化，那么波动率的聚集何以可能？理解这种表面上的矛盾需要我们区分方差和条件方差，不随时间变化的是方差 $\text{Var}(r_t)$ ，聚集的是条件方差 $\text{Var}(r_t | \mathcal{F}_{t-1})$ ，波动率的聚集实际上是一种条件异方差现象。

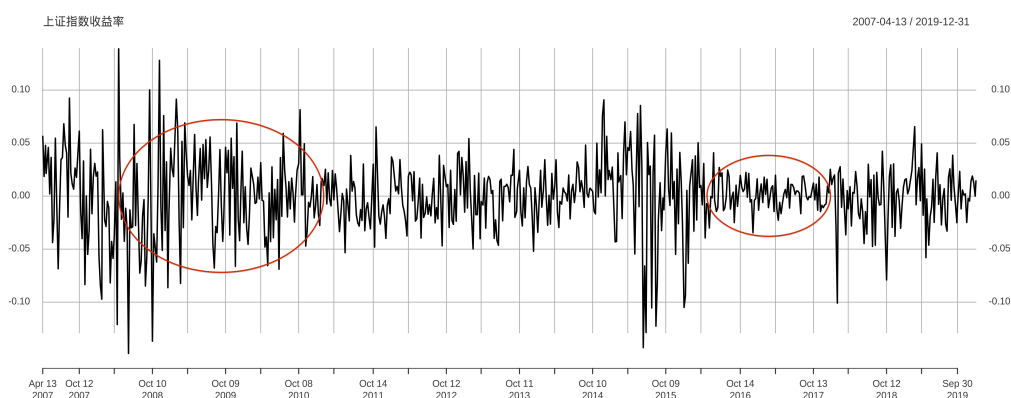


图 1.5: 上证指数对数收益率

为了说明非线性相依性，我们利用如下 R 程序在图 1.6 中绘制了上证指数收益率、收益率的绝对值和收益率的平方序列的 ACF 图³。从图中可以看出，上证指数收益率滞后二十期的自相关系数都不显著。然而，取了绝对值或平方之后的收益率序列呈现出显著的序列相关。

```
1 shindex <- read.table("~/上证指数收益率.csv", sep = ",", header = TRUE)
2 library(forecast)
3 Return <- shindex$rtn
4 absReturn <- abs(Return)
5 Return2 <- Return^2
6
7 acf1 <- ggAcf(Return, 20, main="上证指数收益率")
8 acf2 <- ggAcf(absReturn, 20, main="上证指数收益率绝对值")
9 acf3 <- ggAcf(Return2, 20, main="上证指数收益率平方")
10
11 ggpubr::ggarrange(acf1, acf2, acf3, ncol = 3, nrow = 1)
```

² 例如，图 1.6 中给出的上证指数收益率的 ACF 图表明该收益率在滞后 20 期内不存在明显的序列相关。

³ 关于 ACF 的介绍请参考本书第 2.1 节。

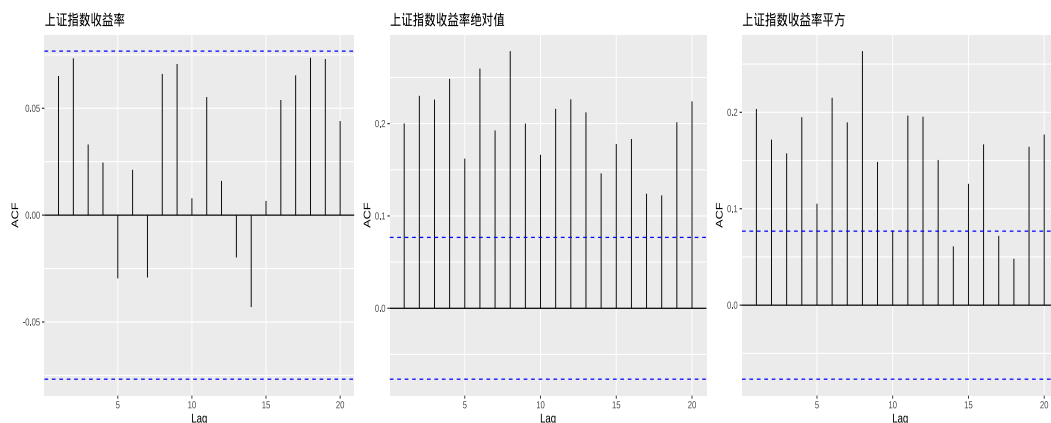


图 1.6: 上证指数对数收益率及其非线性变换的 ACF

1.3 收益率的分布

正态分布常常出现在重要的金融理论的假设中，例如，Markowitz 的资产组合理论和资本资产定价模型 CAPM 都假设了收益率服从正态分布或者投资者偏好是二次的（参阅 Cochrane (2005) 第九章）。虽然对这些理论来说，正态假设并不是必要的（参阅 Ingersoll (1987) 第四章），然而，大量金融文献依然假设正态分布。Mandelbort (1963) 最早对资产收益率服从正态分布的假设提出了根本性的质疑。资产收益率分布的非正态性主要体现在发生极端值的概率大于正态分布预测的概率，这种特征被称为厚尾（heavy tail）或肥尾（fat tail）。此外，资产收益率分布的非对称性也是其非正态特征的体现。

本节讨论如何为收益率的分布建模。考虑在 t 时刻的 N 个资产，它们的收益率为 $\{r_{1t}, \dots, r_{Nt}\}$, $t = 1, 2, \dots, T$ 。关于收益率 $\{r_{it}\}$ 的最一般模型为其联合分布：

$$F(r_{11}, \dots, r_{N1}; r_{12}, \dots, r_{N2}; \dots; r_{1T}, \dots, r_{NT}; \mathbf{x} | \boldsymbol{\theta}),$$

其中， \mathbf{x} 是由状态（state）变量构成的向量，状态变量描述了资产价格形成过程中所处的经济环境； $\boldsymbol{\theta}$ 是唯一确定分布函数 $F(\cdot)$ 的参数向量。金融计量学的任务是在给定 $F(\cdot)$ 和收益率数据 $\{r_{it}\}$ 的情况下对参数 $\boldsymbol{\theta}$ 进行统计推断。不同的资产收益率模型会进一步对 $F(\cdot)$ 施加额外的约束。例如，CAPM 考虑固定时刻 t 之后的收益率横截面 $\{r_{1t}, \dots, r_{Nt}\}$ 的联合分布，这其中假设了收益率的横截面分布不随时间变化。

除了研究收益率的横截面特征之外，金融计量学还关注单个资产收益率的动态特征。考虑资产 i 的收益率序列 $\{r_{i1}, \dots, r_{iT}\}$ 的联合密度 f_i ，我们通常将 f_i 写成如下形式⁴：

$$\begin{aligned} f_i(r_{i1}, \dots, r_{iT}) &= f_{i1}(r_{i1}) \cdot f_{i2}(r_{i2} | r_{i1}) \cdots f_{iT}(r_{iT} | r_{i,T-1}, \dots, r_{i1}) \\ &= f_{i1}(r_{i1}) \prod_{t=2}^T f_{it}(r_{it} | r_{i,t-1}, \dots, r_{i1}). \end{aligned}$$

在上式中，我们需要设定 r_{i1} 的边缘分布 $f_{i1}(\cdot)$ ，接下来我们介绍几种文献中常用的收益率边缘分布。

⁴ 为了符号简洁，我们省略了 $\boldsymbol{\theta}$ 。

正态分布和对数正态分布

早期的金融理论的形成（例如：[Grossman et al. \(1980\)](#) 关于有效市场的讨论）假设简单收益率 $\{R_{it} \mid t = 1, \dots, T\}$ 服从时间维度上的 i.i.d. 正态分布。虽然基于正态假设的模型易于分析和处理，但是假设简单收益率服从正态分布的缺点是显而易见的。首先，持有绝大多数金融资产的责任是有限的，投资者最大的损失为投入的总资金数量。这意味着简单收益率取值的下限是 -1 。然而，服从正态分布的随机变量的实现值是没有下限的；其次，若假设单期简单收益率服从正态分布，则多期简单收益率 $R_{it}[k]$ 不服从正态分布；更重要的是，假设简单收益率服从正态分布与真实收益率数据表现出来的尖峰、厚尾和非对称等典型特征相矛盾。

另一种做法是假设对数收益率 $\{r_{it} \mid t = 1, \dots, T\}$ 服从 i.i.d. 的正态分布。此时，我们称简单收益率服从对数正态分布 (lognormal distribution)。一般地，若 $\log(x) \sim \mathcal{N}(\mu, \sigma^2)$ ，则称 x 服从对数正态分布。因此，对数正态分布的密度函数为：

$$f(x \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} x^{-1} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right), \quad x > 0, \mu \in \mathbb{R}.$$

本章附录 1 给出了对数正态分布的期望和方差：

$$\begin{aligned} \mathbb{E}(x) &= \exp(\mu + \sigma^2/2), \\ \text{Var}(x) &= \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2). \end{aligned}$$

因为 $r_{it} = \log(1 + R_{it})$ ，若 $r_{it} \sim \mathcal{N}(\mu_i, \sigma_i^2)$ ，则有：

$$\begin{aligned} \mathbb{E}(R_{it}) &= \exp(\mu_i + \frac{\sigma_i^2}{2}) - 1, \\ \text{Var}(R_{it}) &= \exp(2\mu_i + \sigma_i^2)[\exp(\sigma_i^2) - 1]. \end{aligned}$$

另一方面，若假设简单收益率 R_{it} 的期望和方差分别为 m_i 和 s_i^2 ，则由上面两个式子可知：

$$\begin{aligned} (m_i + 1)^2 &= \exp(2\mu_i + \sigma_i^2), \\ s_i^2 &= (m_i + 1)^2[\exp(\sigma_i^2) - 1]. \end{aligned}$$

由第二个式子，我们有：

$$\sigma_i^2 = \log \left[1 + \left(\frac{s_i}{m_i + 1} \right)^2 \right].$$

进而可以推出：

$$\mu_i = \log \frac{m_i + 1}{\sqrt{1 + \left(\frac{s_i}{m_i + 1} \right)^2}}.$$

对数正态分布是非对称的右侧厚尾分布，图 1.7 绘制了 $\mu = 0$ ， σ^2 分别为 $1/16$ 、 $1/4$ 和 1 的三条对数正态分布的概率密度函数曲线。从图中可以看出， σ^2 越大，对数正态分布的尾部越厚。假设对数收益率服从正态分布有以下优点：

- 多期对数收益率 $r_t[k]$ 依然服从正态分布；
- 对数收益率 r_t 没有下界限制；

- $1 + R_t = \exp(r_t)$ 非负，符合有限责任制度。

然而，假设对数收益率服从正态分布与真实收益率数据表现出来的典型特征也是不一致的。

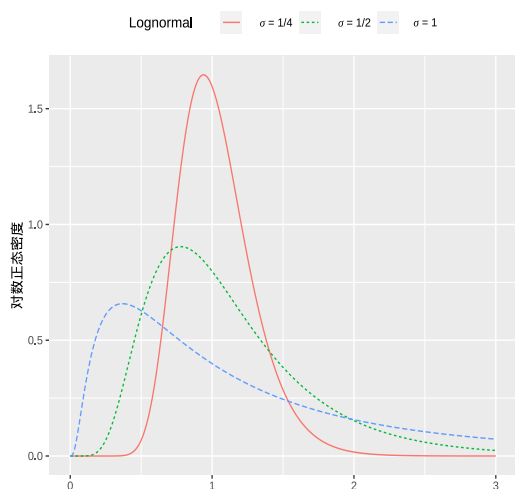


图 1.7: 对数正态分布的概率密度函数

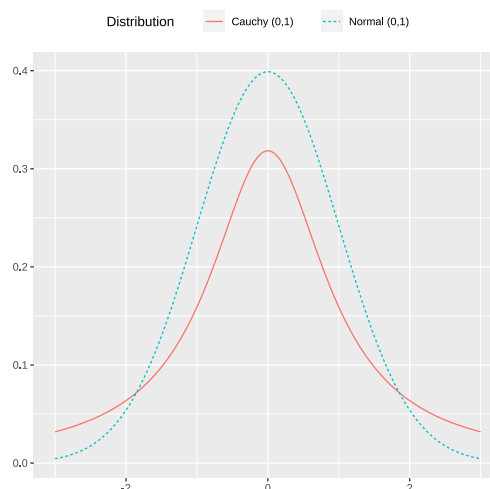


图 1.8: Cauchy 分布和标准正态分布

为了体现收益率分布的非对称性，金融时间序列分析中有时会使用有偏的正态分布，其概率密度函数为：

$$f(x | \mu, \sigma, \alpha) = \frac{2}{\sigma} \cdot \phi\left(\frac{x - \mu}{\sigma}\right) \cdot \Phi\left(\alpha \cdot \frac{x - \mu}{\sigma}\right),$$

其中， μ 是位置参数； σ 是尺度 (scale) 参数； α 为偏度参数； $\phi(\cdot)$ 为标准正态分布的 pdf； $\Phi(\cdot)$ 为标准正态分布的 cdf。图 1.9 绘制了 ($\mu = 0, \sigma = 1$) 以及 α 分别取 1, 3 和 -3 的三条有偏正态分布的概率密度函数曲线，并与标准正态分布概率密度函数曲线进行对比。从图中可以看出，当 $\alpha < 0$ 时，有偏正态分布负偏；当 $\alpha > 0$ 时，有偏正态分布正偏。

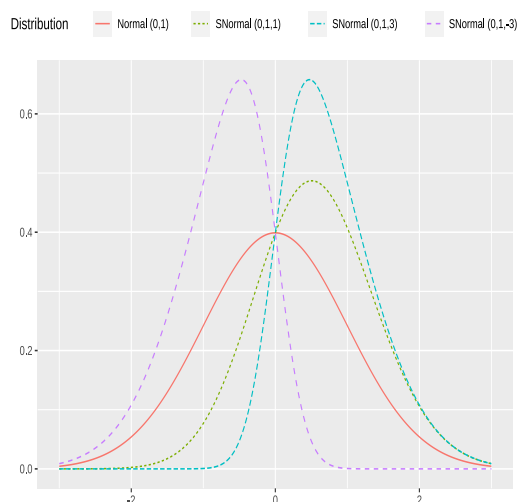


图 1.9: 有偏正态分布的密度函数

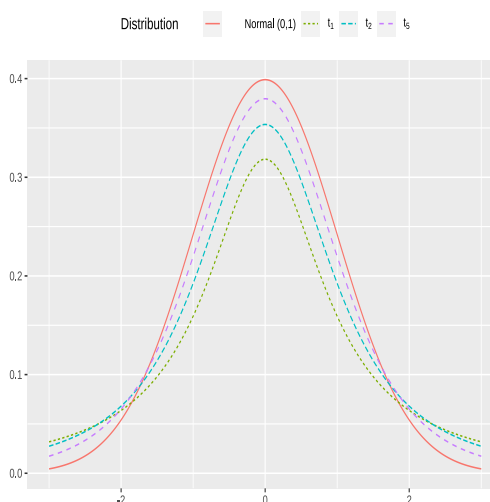


图 1.10: 学生 t 分布的密度函数

稳定分布

鉴于资产收益率表现出的厚尾特征，Mandelbort (1963) 推测使用非正态的稳定分布 (stable distribution) 为收益率建模应该是更恰当的。稳定分布是正态分布的自然推广，此类分布对加法保持稳定，即，服从稳定分布的随机变量之和仍然服从稳定分布。为了区分正态和非正态的稳定分布，后者常被称为“稳定帕累托”分布或者“Lévy 稳定”分布。Fama (1963) 进一步巩固了资产收益率更适合用稳定帕累托分布建模的观点。Cauchy 分布是具有解析密度函数表达式的稳定帕累托分布之一，其概率密度函数为：

$$f(x | \gamma, \delta) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - \delta)^2},$$

其中， δ 为位置参数， γ 为尺度参数。

然而，真实收益率分布的尾部厚度往往小于稳定帕累托分布的尾部厚度。事实上，稳定帕累托分布的尾部非常厚以致于此类分布的方差和更高阶矩都是无穷大的。图 1.8 绘制了标准正态分布和 ($\delta = 0, \gamma = 1$) 的 Cauchy 分布的概率密度函数曲线，从图中可以看出，Cauchy 分布的尾部厚于标准正态分布的尾部。因为标准的金融理论一般都要求收益率具有有限的二阶（或更高阶）矩，所以稳定帕累托分布如今已经很少被用来刻画资产收益率。

此外，利用服从稳定帕累托分布的随机变量的一组样本估计的期望和标准差并不随着样本量的增加而收敛，然而利用真实收益率数据估计期望和方差时估计量往往收敛。同时，若假设对数收益率服从稳定帕累托分布，则多期对数收益率依然服从稳定帕累托分布。但是，通过对真实的资产收益率数据进行的分析发现多期对数收益率的分布会随着持有资产期数的增加而越来越接近正态。

关于稳定帕累托分布的详细讨论可以参考 Rachev et al. (2005) 的第七章。

学生 t 分布和有偏的学生 t 分布

（自由度大于 2 的）学生 t 分布经常被用来刻画收益率的分布，此类分布的二阶矩有限且具有厚尾特征。自由度为 r 的 t 分布的概率密度函数为：

$$f(x | r) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi} \cdot \Gamma\left(\frac{r}{2}\right)} \left(1 + \frac{x^2}{r}\right)^{-\left(\frac{r+1}{2}\right)}, \quad -\infty < x < \infty,$$

其中， $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ 为 Gamma 函数。注意到 $\Gamma(1) = 1$, $\Gamma(1/2) = \sqrt{\pi}$ ，因此，当 $r = 1$ 时， t 分布即为 Cauchy 分布。此外，可以证明：当 $r \rightarrow \infty$ 时， t 分布依分布收敛到标准正态分布。对于自由度 $r > 1$ 的 t 分布，其期望为 0；对于自由度 $r > 2$ 的 t 分布，其方差为 $r/(r-2)$ ；一般地，对于自由度为 r 的 t 分布，其前 $r-1$ 阶矩是存在的。图 1.10 分别绘制了自由度为 1, 2 和 5 的三条 t 分布的密度函数曲线以及标准正态分布的密度函数曲线，从图中可以看出，随着自由度的增加， t 分布的曲线逐渐接近标准正态分布的曲线。

在第 10 章要讨论的波动率模型中（如：ARCH 和 GARCH），我们经常需要将随机扰

动项的方差标准化成 1，为此，引入如下 scaled- t 分布的密度函数：

$$f(x | r, v) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi v} \cdot \Gamma\left(\frac{r}{2}\right)} \left(1 + \frac{x^2}{rv}\right)^{-\left(\frac{r+1}{2}\right)}, \quad -\infty < x < \infty,$$

其中， v 是尺度参数。上述 scaled- t 分布的方差为 $vr/(r-2)$ 。因此，为了使方差为 1，可以取 $v = (r-2)/r$ 。

为了同时刻画收益率分布的厚尾和有偏两个特征，Fernández et al. (1998) 提出了将关于 0 对称并且单峰（unimodal）的分布转换为有偏分布的一般方法。具体地，令 $f(\cdot)$ 为某一对称单峰分布的概率密度函数，给定偏度参数 $\gamma \in (0, \infty)$ ，基于 f 的有偏分布的 pdf 可以按如下方式构造：

$$p(z | \gamma) = \frac{2}{\gamma + \gamma^{-1}} \left\{ f\left(\frac{z}{\gamma}\right) \mathbf{1}_{[0, \infty)}(z) + f(\gamma z) \mathbf{1}_{(-\infty, 0)}(z) \right\},$$

其中，

$$\mathbf{1}_A(z) = \begin{cases} 1 & z \in A, \\ 0 & z \notin A. \end{cases}$$

显然，当 $\gamma \neq 1$ 时， $p(z | \gamma)$ 是非对称的。事实上，我们有：

$$p(z | \gamma = 1) = f(z),$$

以及

$$\frac{\mathbb{P}(z \geq 0 | \gamma)}{\mathbb{P}(z < 0 | \gamma)} = \gamma^2.$$

从上式可以看出， γ 控制了随机变量 z 的实现值落入 0 两侧的几率比。可以证明：

$$\mathbb{E}(z^m | \gamma) = M_m \frac{\gamma^{m+1} + (-1)^m \gamma^{-(m+1)}}{\gamma + \gamma^{-1}}, \quad m \in \mathbb{N}^+,$$

其中，

$$M_m = 2 \int_0^\infty s^m f(s) ds.$$

特别地，

$$\mathbb{E}(z) = M_1 (\gamma - \gamma^{-1}),$$

$$\text{Var}(z) = (M_2 - M_1^2) (\gamma^2 + \gamma^{-2}) + 2M_1^2 - M_2.$$

可以利用 $\mathbb{E}(z)$ 和 $\text{Var}(z)$ 将有偏分布标准化为期望为 0 且方差为 1 的概率分布。

混合正态分布

混合（mixture）正态分布的密度函数为：

$$f(x) = \sum_{k=1}^K \pi_k \cdot \phi(x | \mu_k, \sigma_k^2), \quad \sum_{k=1}^K \pi_k = 1$$

其中， $\phi(\cdot | \mu, \sigma^2)$ 是期望为 μ ，方差为 σ^2 的正态分布的 pdf； π_k 是第 k 个正态分布在混合正态中的权重。混合正态分布提供了由多个简单概率分布构造出复杂概率分布的一般

框架。图 1.11 分别绘制了标准正态分布、 $\mathcal{N}(0, 0.25)$ 、 $\mathcal{N}(0, 25)$ 以及混合正态分布：

$$0.5 \times \mathcal{N}(0, 1) + 0.4 \times \mathcal{N}(0, 0.25) + 0.1 \times \mathcal{N}(0, 25)$$

四条概率密度曲线，从图中可以看出，三个正态分布混合之后得到的概率分布具有厚尾的特征。不仅如此，混合正态分布形式非常灵活，可以具有有偏、双峰或多峰等特征。例如，考虑如下混合正态分布⁵：

- 单峰有偏分布：

$$\frac{1}{5} \times \mathcal{N}(0, 1) + \frac{1}{5} \times \mathcal{N}\left(\frac{1}{2}, \left(\frac{2}{3}\right)^2\right) + \frac{3}{5} \times \mathcal{N}\left(\frac{13}{12}, \left(\frac{5}{9}\right)^2\right);$$

- 双峰有偏分布：

$$\frac{3}{4} \times \mathcal{N}(0, 1) + \frac{1}{4} \times \mathcal{N}\left(\frac{3}{2}, \left(\frac{1}{3}\right)^2\right);$$

- 三峰分布：

$$\frac{9}{20} \times \mathcal{N}\left(-\frac{6}{5}, \left(\frac{3}{5}\right)^2\right) + \frac{9}{20} \times \mathcal{N}\left(\frac{6}{5}, \left(\frac{3}{5}\right)^2\right) + \frac{1}{10} \times \mathcal{N}\left(0, \left(\frac{1}{4}\right)^2\right).$$

图 1.12 绘制了上述三个混合正态分布的概率密度函数曲线。

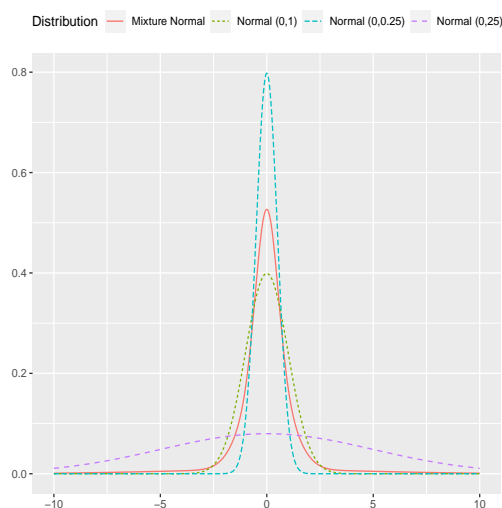


图 1.11: 正态分布和混合正态分布（对称）

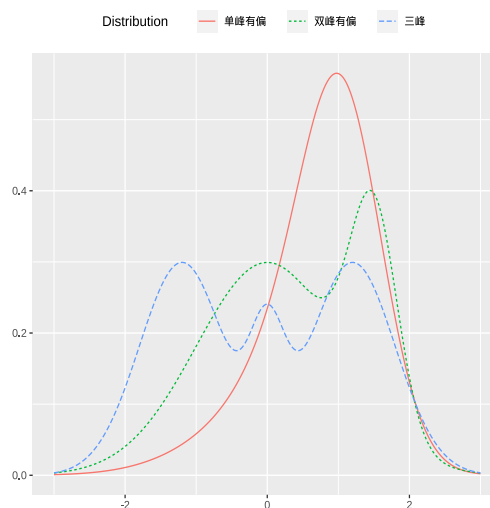


图 1.12: 单峰和多峰混合正态分布

虽然混合正态分布继承了正态分布的易处理、各阶矩有限等特点，并且可以刻画分布的厚尾和非对称等特征，但是混合权重的估计比较困难，本书附录 A 将要介绍估计混合权重的 EM 算法，关于混合正态更详细的介绍可以参考 Bishop (2006)。

1.4 有效市场假说和收益率的可预测性 *

有效市场假说 (Efficient Markets Hypothesis, EMH) 是整个金融学研究体系的基础性理论之一，它的历史由来已久，至少可以追溯到 Bachelier (1900)。Fama (1970) 将有效市场定义为 “A market in which prices always "fully reflect" available information is called

⁵ 更多关于混合正态分布的例子请参考 Marron et al. (1992)。

"efficient." (资产价格充分反映了可获取的所有信息的市场)。EMH 的基本逻辑是：若价格是可以预测的，则交易者将具有获取超常 (superior) 收益的机会，大量追求利益的交易者之间的竞争会迅速消灭这些机会。因此，EMH 蕴涵着价格的不可预测性。对这一逻辑略带讽刺意味的经典笑话是：两位芝加哥学派的经济学家在人行道上发现 10 美元纸币，其中一位阻止另一位捡起这 10 美元，并说：“这是不可能的，否则它早就被别人捡走了”。⁶

根据 Fama 的定义，市场的有效性强弱与信息集有关。按照信息集的不同，市场的有效性可以分为以下三类：

- **弱式有效**：信息集中只包括历史价格或收益率。当前价格充分反映了历史价格的信息，仅基于历史价格信息构造的交易策略无法获得异常 (abnormal) 收益；
- **半 (Semi-) 强式有效**：信息集中包括所有市场参与者都可以获取的信息，即，公开可得信息，包括：历史价格、年报、盈利预期和宏观指标等。当前价格充分反映了公开可得的信息，仅基于公开可得信息构造的交易策略无法获得异常收益；
- **强式有效**：信息集中包括任一市场参与者能获取的所有信息，即，公开可得信息和私有信息。当前价格充分反映了公开可得的信息和私有信息，基于公开可得信息和私有信息构造的交易策略无法获得异常收益。

为了计算异常收益，我们需要设定一个资产定价模型用来定义风险溢价 (risk premium)，进而定义资产的“正常”收益。在此基础上，资产的异常收益定义为资产收益与它的正常收益之差。在给定信息集的情况下，若资产的异常收益是不可预测的、具有“随机”性，则 EMH 不能被拒绝。

令 r_t^* 是基于某一资产定价模型计算的在时间区间 $[t, T]$ 上持有资产获得的正常收益，这里假设 $T > t$ 。在 t 时刻， r_t^* 是已知的。EMH 要求任意风险资产在相同时间区间上的现实收益 r_T 满足：

$$\mathbb{E}[r_T - r_t^* \mid \mathcal{F}_t] = 0, \quad (1.1)$$

其中， \mathcal{F}_t 表示 t 时刻的信息集，例如，对于弱式有效， \mathcal{F}_t 包含了 t 时刻之前（包括 t 时刻）的风险资产的历史价格。上式表明，在平均意义上，投资者不可能获取比 r_t^* 更高的收益。由定理 1.2，强式有效可以推出半强式有效，半强式有效可以推出弱式有效。事实上，令 $\mathcal{F}_t' \subseteq \mathcal{F}_t$ ，根据迭代期望率，我们有：若式 (1.1) 成立，

$$\mathbb{E}[r_T - r_t^* \mid \mathcal{F}_t'] = \mathbb{E}\{\mathbb{E}[r_T - r_t^* \mid \mathcal{F}_t] \mid \mathcal{F}_t'\} = 0.$$

此外，由定理 1.5，

$$\text{Var}(r_T \mid \mathcal{F}_t') = \mathbb{E}[\text{Var}(r_T \mid \mathcal{F}_t) \mid \mathcal{F}_t'] + \text{Var}[\mathbb{E}(r_T \mid \mathcal{F}_t) \mid \mathcal{F}_t'].$$

对上式两端取期望得到：

$$\mathbb{E}[\text{Var}(r_T \mid \mathcal{F}_t')] \geq \mathbb{E}[\text{Var}(r_T \mid \mathcal{F}_t)], \quad (1.2)$$

⁶ 反对 EMH 的主要金融理论是行为金融学，对此方向感兴趣的读者可以参考李勉群 (美)，苏子英 (美)【著】，张然【译】(2019)，这里就不再赘述了。

即，在平均意义上，增加信息可以减小条件方差。同时，由定理 1.4，

$$\begin{aligned}\text{Var}(r_T) &= \mathbb{E}\{\text{Var}[r_T | \mathcal{F}_t]\} + \text{Var}\{\mathbb{E}[r_T | \mathcal{F}_t]\}, \\ &= \mathbb{E}\{\text{Var}[r_T | \mathcal{F}'_t]\} + \text{Var}\{\mathbb{E}[r_T | \mathcal{F}'_t]\}.\end{aligned}$$

根据式 (1.2) 可知： $\text{Var}\{\mathbb{E}[r_T | \mathcal{F}_t]\} \geq \text{Var}\{\mathbb{E}[r_T | \mathcal{F}'_t]\}$ ，即，随着信息的增加，预测的波动加剧。

Shleifer et al. (1997) 和 Grossman et al. (1980) 分别从套利限制和信息收集成本两个角度对 EMH 提出了质疑，推荐感兴趣的读者阅读这些金融经济学领域的经典文献。本节主要讨论如何对 EMH 进行检验，这里有几个问题需要说明：(1) 计算正常收益需要设定资产定价模型，如果拒绝 EMH 成立的原假设，我们无法判断是市场无效还是正常收益的模型误设，这被称为联合假设问题 (joint hypothesis problem)；(2) 信息集的选择与假设检验的结果有关，基于小的信息集的检验无法拒绝 EMH 的原因可能是因为选取了没有信息量的信息集；(3) 在构造统计检验时，应该保留哪些关于收益率性质的假设也是值得考虑的问题。

鞅

令 P_t 是某一资产在 t 时刻的价格， $\{P_t\}$ 是鞅意味着：

$$\mathbb{E}[P_{t+1} | P_t, P_{t-1}, \dots] = P_t,$$

或等价地写成：

$$\mathbb{E}[P_{t+1} - P_t | P_t, P_{t-1}, \dots] = 0.$$

鞅假设表明：给定资产价格的历史信息，资产今天的价格是其明天价格的最好预测。鞅长期以来被认为是有效市场的必要条件。事实上，若市场是有效的，通过交易资产历史价格中所包含的信息从而获利是不可能的。以历史价格为条件，未来价格变化的条件期望一定是 0。在一个完全有效的市场中，价格的变化是完全随机的、不可预测的。

通常，价格被认为是由未来现金流的折现决定的，如果事实如此，那么价格变化的随机性似乎难以理解。然而，价格的现值模型和价格变化的随机性是不存在矛盾的。令 D_t 表示 t 时刻的现金流， R 表示折现率，现值模型指出：⁷

$$P_t = \sum_{i=1}^{\infty} \left(\frac{1}{1+R} \right)^i \mathbb{E}_t(D_{t+i}),$$

其中， $\mathbb{E}_t(\cdot) = \mathbb{E}(\cdot | \mathcal{F}_t)$ 。令 $V^* = \sum_{i=1}^{\infty} \left(\frac{1}{1+R} \right)^i D_{t+i}$ ，则有：

$$P_t = \mathbb{E}_t(V^*).$$

同理，

$$P_{t+1} = \mathbb{E}_{t+1}(V^*).$$

由定理 1.5，我们有：

$$\mathbb{E}[P_{t+1} - P_t] = \mathbb{E}_t[\mathbb{E}_{t+1}(V^*) - \mathbb{E}_t(V^*)] = 0.$$

⁷ 见 Linton (2019) 第九章的式 (9.3)。

因此，给定信息集 \mathcal{F}_t ，价格变化依然是不可预测的。

鞅和随机游走模型有着密切的联系，随机游走模型是检验 EMH 时常用的模型。

随机游走

随机游走模型常用来刻画随机过程的不可预测性。考虑离散时间的随机过程 P_t , $t = 1, 2, \dots$ ，若 P_t 满足：

$$P_t = P_{t-1} + \varepsilon_t,$$

其中， ε_t 是期望为 0 的冲击项，则称 $\{P_t\}$ 为随机游走过程或酒鬼漫步过程，这是因为，给定酒鬼的当前位置，下一时刻他的位置是难以预测的。下面的 R 程序模拟了一个随机游走过程，并绘制了该过程的时间序列图和 ACF 图，其输出结果见图 1.13。

```
1 library(tidyverse)
2 library(forecast)
3 set.seed(20220719)
4 e <- rnorm(300) # 生成300个标准正态随机数
5 x <- e # 创建300*1向量x用以保存模拟数据
6 for(i in 2:300) x[i] <- x[i-1] + e[i] # 模拟随机游走过程
7 x <- as.data.frame(cbind(c(1:300),x))
8 names(x) <- c("time","position")
9 rwplot <- x %>%
10   ggplot(aes(time,position)) +
11   geom_line()
12 rwacf <- ggAcf(x$position,50,main="")
13 ggpubr::ggarrange(rwplot,rwacf, ncol = 2, nrow = 1)
```

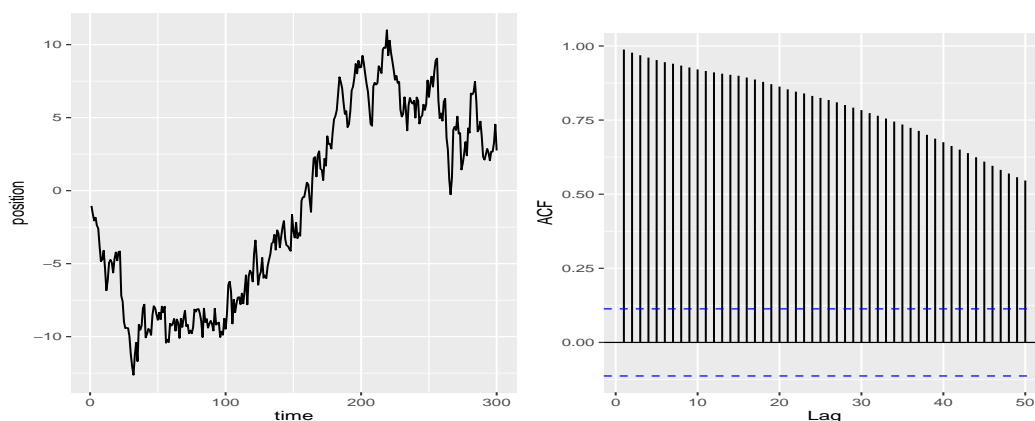


图 1.13: 无漂移项的随机游走过程及其 ACF

具有漂移项的随机游走过程满足：

$$P_t = \mu + P_{t-1} + \varepsilon_t,$$

其中，漂移项 μ 可以是不为 0 的常数。将输出图 1.13 结果的 R 程序的第 6 行改为：

```
1 for(i in 2:300) x[i] <- 0.1 + x[i-1] + e[i] # 模拟随机游走过程
```

我们得到如下漂移项 $\mu = 0.1$ 的随机游走的时间序列图和 ACF 图：

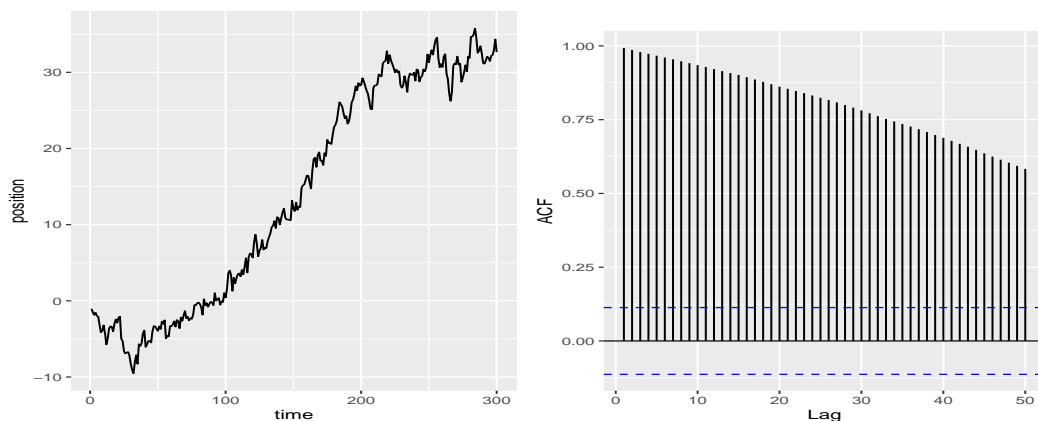


图 1.14: 具有漂移项 $\mu = 0.1$ 的随机游走过程及其 ACF

通过对比图 1.13 和图 1.14，我们发现两种随机游走的 ACF 的特征是相似的，在小的滞后期时，ACF 接近于 1，随着滞后期的增加，ACF 逐渐衰减，但是衰减的非常缓慢。但是，两种随机游走的时间序列图存在着明显的差异， $\mu = 0.1$ 的随机游走具有一个向上的时间趋势，而无漂移项的随机游走没有这种趋势。为了理解时间趋势的由来，设随机游走的初值为 P_0 ，则有：

$$\begin{aligned} P_1 &= \mu + P_0 + \varepsilon_1 \\ P_2 &= \mu + P_1 + \varepsilon_2 = 2\mu + P_0 + \varepsilon_2 + \varepsilon_1 \\ &\vdots \\ P_t &= t\mu + P_0 + \varepsilon_t + \cdots + \varepsilon_1. \end{aligned}$$

因此，具有漂移项的随机游走在 t 时刻的位置由时间趋势 $t\mu$ 、初始位置 P_0 和累积的随机冲击 $\sum_{s=1}^t \varepsilon_s$ 三部分组成。

对于金融时间序列来说， μ 表示资产的正常收益。一般地， μ 是由描述风险溢价动态演变过程的特定资产定价模型计算得出的，因此， μ 具有时变（time-varying）性，可以记为 μ_t 。当数据的频率相对较高时（例如：日频或周频）， μ_t 的变化通常可以忽略不计，为了讨论问题的方便，我们假设 μ_t 为常数。

精确设定随机游走模型还需要给出关于 ε_t 的明确假设。考虑如下几种关于 ε_t 的不同假设：

- RW1: ε_t 是 i.i.d., 且满足 $\mathbb{E}(\varepsilon_t) = 0$;
- RW2: 不同时刻的 ε_t 彼此独立, 且满足 $\mathbb{E}(\varepsilon_t) = 0$;
- RW3: $\{\varepsilon_t\}$ 是鞅差序列, 即, 对任意的 t , $\mathbb{E}(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = 0$ 以概率 1 成立;
- RW4: 对任意的 $k > 0$ 和 t , $\mathbb{E}(\varepsilon_t | \varepsilon_{t-k}) = 0$ 以概率 1 成立;
- RW5: 对任意的 $k > 0$ 和 t , $\mathbb{E}(\varepsilon_t) = 0$, $\text{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$ 。

若 $\mathbb{E}(\varepsilon_t^2) < \infty$, 根据定理 1.5, 我们有: $\text{RW1} \Rightarrow \text{RW2} \Rightarrow \text{RW3} \Rightarrow \text{RW4} \Rightarrow \text{RW5}$ 。RW3-5 排除了 ε_t 在时间维度上的相关性, 即, $\text{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0, k > 0$; 但是, 这三个假设允许 ε_t 在时间维度上存在相依 (dependent) 关系。例如, $\text{Cov}(\varepsilon_t^2, \varepsilon_{t-k}^2) \neq 0$ 并不违背这三条假设。金融计量学的经典教材 **Campbell et al. (1997)** 讨论了上述假设中的 RW1、RW2 和 RW5, 以及如何检验这三类假设。然而, **Fama (1970)** 认为最合适的假设应该是 RW3。在 RW3 成立的情况下,

$$\text{Cov}[\varepsilon_t, g(P_{t-1}, P_{t-2}, \dots)] = 0,$$

其中, $g(\cdot)$ 是满足相关矩条件的任意 (可测) 函数。从上式可以进一步看出 RW3 要弱于 RW2, 因为 RW2 意味着: 对任意的可测函数 $h(\cdot)$ 和 $g(\cdot)$,

$$\text{Cov}[h(\varepsilon_t), g(P_{t-1}, P_{t-2}, \dots)] = 0.$$

随机游走检验: RW1

本小节讨论随机游走的假设检验。检验统计量的构造依赖于我们保留的关于 ε_t 的假设, 通常, 假设条件越强, 检验统计量的形式越简单。

关于 RW1 的最早的检验之一是由 **Cowles et al. (1937)** 提出的同向/反向 (Sequences and Reversals) 比检验。假设对数价格满足:

$$p_t = p_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim i.i.d. (0, \sigma^2).$$

定义对数价格上涨的示性函数 I_t 如下:

$$I_t = \begin{cases} 1, & \text{若 } r_t \equiv p_t - p_{t-1} > 0, \\ 0, & \text{若 } r_t \equiv p_t - p_{t-1} \leq 0. \end{cases}$$

Cowles et al. (1937) 的同向/反向比检验的想法是比较收益率的历史数据中出现同向变化和反向变化的频率。具体地, 定义同向变化为:

$$r_t \times r_{t+1} > 0 \quad \text{或者} \quad \text{sign}(r_t) = \text{sign}(r_{t+1}).$$

同理, 定义反向变化为:

$$r_t \times r_{t+1} < 0 \quad \text{或者} \quad \text{sign}(r_t) = -\text{sign}(r_{t+1}).$$

基于上述定义, 可以构造如下 Cowles-Jones 比率:

定义 1.1. Cowles-Jones 比率

Cowles-Jones 比率定义为:

$$CJ = \frac{\mathbb{P}(\text{同向变化})}{\mathbb{P}(\text{反向变化})} = \frac{\mathbb{P}(r_t \times r_{t+1} > 0)}{1 - \mathbb{P}(r_t \times r_{t+1} > 0)}.$$



假设对数收益率 r_t 在时间维度上相互独立 (RW2), 并且 $\pi = \mathbb{P}(r_t > 0)$ 为常数, 则有:

$$\begin{aligned} \pi_s &= \mathbb{P}(\text{同向变化}) = \mathbb{P}(r_t \times r_{t+1} > 0) \\ &= \mathbb{P}(r_t > 0, r_{t+1} > 0) + \mathbb{P}(r_t < 0, r_{t+1} < 0) \end{aligned}$$

$$\begin{aligned}
&= \mathbb{P}(r_t > 0)\mathbb{P}(r_{t+1} > 0) + \mathbb{P}(r_t < 0)\mathbb{P}(r_{t+1} < 0) \\
&= \pi^2 + (1 - \pi)^2.
\end{aligned}$$

Cowles et al. (1937) 假设 $\pi = 0.5$, 从而 $\pi_s = 0.5$, $CJ = 1$ 。我们这里采用 Linton (2019) 的处理方式, 允许 $\pi \neq 0.5$ 。实际上, 若对数价格满足具有漂移项的随机游走, 即,

$$p_t = \mu + p_{t-1} + \varepsilon_t,$$

则 $\pi = \mathbb{P}(r_t > 0)$ 将不等于 0.5。例如, 假设 $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$, 则有:

$$I_t = \begin{cases} 1, & \text{概率为 } \pi, \\ 0, & \text{概率为 } 1 - \pi, \end{cases}$$

其中, $\pi = 1 - \Phi\left(\frac{-\mu}{\sigma}\right)$ 。因此, 若 $\mu > 0$, 则 $\pi > 0.5$; 若 $\mu < 0$, 则 $\pi < 0.5$ 。并且,

$$CJ = \frac{\pi_s}{1 - \pi_s} = \frac{\pi^2 + (1 - \pi)^2}{2\pi(1 - \pi)} \geq 1.$$

于是, 只要漂移项不为 0, 同向变化发生的概率就会大于反向变化发生的概率, 这是因为非零的漂移项在随机游走过程中引入了趋势。

应用 CJ 进行随机游走检验时需要对其进行估计, 为此定义如下经验 CJ 比率:

定义 1.2. 经验 Cowles-Jones 比率

令 $\{r_1, \dots, r_{T+1}\}$ 为资产对数收益率的一组样本观测, T_s 为同向变化的次数, 经验 Cowles-Jones 比率定义为:

$$\widehat{CJ} = \frac{T_s}{T - T_s} = \frac{T_s/T}{1 - T_s/T},$$

其中, $T_s = \sum_{t=1}^T D_t$, $D_t = I_t I_{t+1} + (1 - I_t)(1 - I_{t+1})$ 。



\widehat{CJ} 的渐近分布可以利用 Delta 方法求得。具体地, 注意到 T_s 可以看作 T 个 Bernoulli 随机变量 D_t 之和, 并且

$$D_t = \begin{cases} 1, & \text{概率为 } \pi_s, \\ 0, & \text{概率为 } 1 - \pi_s. \end{cases}$$

因此, $\mathbb{E}(T_s) = T\pi_s$ 。同时, 在 RW1 条件下, 相邻两期 D_t 存在相关性, 我们有:

$$\begin{aligned}
\text{Var}[T_s] &= T\pi_s(1 - \pi_s) + 2T\text{Cov}[D_t, D_{t+1}] \\
&= T\pi_s(1 - \pi_s) + 2T[\pi^3 + (1 - \pi)^3 - \pi_s^2].
\end{aligned}$$

由中心极限定理, 我们有:

$$\sqrt{T} \left(\frac{T_s}{T} - \pi_s \right) \xrightarrow{d} \mathcal{N}(0, V_s),$$

其中, $V_s = \pi_s(1 - \pi_s) + 2[\pi^3 + (1 - \pi)^3 - \pi_s^2]$ 。再根据 Delta 方法,

$$\sqrt{T} \left(\widehat{CJ} - \frac{\pi_s}{1 - \pi_s} \right) \xrightarrow{d} \mathcal{N}(0, V_{CJ}),$$

其中, $V_{CJ} = V_s/(1 - \pi_s)^4$ 。为了使用上述结论构造检验统计量, 我们需要估计 π 进而得到 π_s 和 V_{CJ} 的估计。因为, $\pi = \mathbb{P}(r_t > 0)$, 所以, π 的一个直观的估计量是正的收益率在样本中出现的频率, 记为 $\hat{\pi}$ 。于是, 我们有如下结论:

定理 1.2. 同向/反向比检验

若 RW1 成立，并且 $\pi = \mathbb{P}(r_t > 0)$ 为常数，则有：

$$S_{CJ} \equiv \sqrt{T} \left(\frac{\widehat{CJ} - \hat{\pi}_s / (1 - \hat{\pi}_s)}{\sqrt{\widehat{V}_{CJ}}} \right) \xrightarrow{d} \mathcal{N}(0, 1),$$

其中， $\hat{\pi}_s = \hat{\pi}^2 + (1 - \hat{\pi})^2$ ，类似地，可以计算 \widehat{V}_{CJ} 。



若 $|S_{CJ}| > z_{1-\frac{\alpha}{2}}$ ，我们可以拒绝 RW1 成立的原假设。

检验 RW1 的另一种常用方法是游程 (runs) 检验。在一个只含有 0 和 1 的字符串中，连续的 1 序列或者连续的 0 序列被称为一个游程。对于对数收益率的时间序列，我们可以利用 I_t 将其编码为由 0 和 1 构成的字符串，例如，10 个对数收益率构成的时间序列被编码为 1001110100，此字符串的游程个数是 6，包括 3 个由 1 构成的正游程 (1, 111, 1) 和 3 个由 0 构成的负 (对应着资产价格下跌) 游程 (00, 0, 00)。游程检验的基本思想是通过比较数据中的实际游程数和 RW1 假设下的期望游程数构造检验统计量。

记 $N_{\text{runs}}(1)$ 为样本中的正游程数， $N_{\text{runs}}(0)$ 为样本中的负游程数， $N_{\text{runs}} = N_{\text{runs}}(1) + N_{\text{runs}}(0)$ 为总游程数。在 RW1 成立的情况下，Campbell et al. (1997) 给出了如下检验统计量及其渐近分布：

$$z = \frac{N_{\text{runs}} - 2T\hat{\pi}(1 - \hat{\pi})}{2\sqrt{T\hat{\pi}(1 - \hat{\pi})[1 - 3\hat{\pi}(1 - \hat{\pi})]}} \xrightarrow{d} \mathcal{N}(0, 1),$$

其中， T 为样本量， $\hat{\pi}$ 表示正的收益率在样本中出现的频率。在实际使用中，往往对上式的 z 统计量做如下的连续修正：

$$z = \frac{N_{\text{runs}} + \frac{1}{2} - 2T\hat{\pi}(1 - \hat{\pi})}{2\sqrt{T\hat{\pi}(1 - \hat{\pi})[1 - 3\hat{\pi}(1 - \hat{\pi})]}} \xrightarrow{d} \mathcal{N}(0, 1).$$

这是因为：在区间 $[N_{\text{runs}}, N_{\text{runs}} + 1)$ 上， z 统计量的实现值不同从而导致正态近似下的 p 值也不同。然而， N_{runs} 是整数，在区间 $[N_{\text{runs}}, N_{\text{runs}} + 1)$ 上的 p 值应该是常数。

随机游走检验：RW5

假设对数价格满足：

$$p_t = \mu + p_{t-1} + \varepsilon_t,$$

其中， ε_t 满足 RW5。这意味着对数收益率满足：

$$r_t = \mu + \varepsilon_t.$$

显然，若 ε_t 不存在自相关，则对数收益率也不存在自相关。因此，检验 RW5 可以通过 r_t 的自相关检验实现。可观测时间序列的自相关检验方法已在第 3.2 节中做过详细地介绍，在此不再赘述。

随机游走检验：方差比检验

若 RW1-RW5 中的任意一个条件成立，则多期对数收益率的方差具有线性性质。例如，若 RW1 成立，两期对数收益率的方差为 $\text{Var}(r_t[2]) = \text{Var}(r_t + r_{t-1}) = 2\text{Var}(r_t)$ ；若

RW5 成立，两期对数收益率的方差依然可以表示为 $\text{Var}(r_t[2]) = \text{Var}(r_t) + \text{Var}(r_{t-1})$ 。⁸

为了讨论问题的方便，假设对数收益率序列 $\{r_t\}$ 是弱平稳时间序列。定义方差比 (Variance Ratio) 为：

$$VR(q) = \frac{\text{Var}(r_t[q])}{q\text{Var}(r_t)}.$$

在 RW5 条件下，容易验证 $VR(2) = 1$ 。一般情况下， $\{r_t\}$ 可能存在自相关，则有：

$$VR(2) = \frac{2\text{Var}(r_t) + 2\text{Cov}(r_t, r_{t-1})}{2\text{Var}(r_t)} = 1 + \rho_1.$$

当且仅当 $\rho_1 = 0$ 时， $VR(2) = 1$ 。类似地，对于一般的多期对数收益率 $r_t[q]$ ，

$$\begin{aligned} VR(q) &= \frac{\text{Var}(r_t[q])}{q\text{Var}(r_t)} = \frac{q\text{Var}(r_t) + 2\sum_{k=1}^{q-1}(q-k)\text{Cov}(r_t, r_{t-k})}{q\text{Var}(r_t)} \\ &= 1 + 2\sum_{k=1}^{q-1}\left(1 - \frac{k}{q}\right)\rho_k. \end{aligned} \quad (1.3)$$

若 RW1-RW5 中的任意一个条件成立，对任意的 $q > 1$ ，我们都有 $VR(q) = 1$ 。方差比检验的基本思想就是比较方差比的样本估计值和 1 之间的差别从而判断数据是否满足随机游走假设。

另一方面，当随机游走假设不成立时，式 (1.3) 也可以提供关于备择假设的额外信息，原因在于式 (1.3) 体现了在备择假设下时间序列的自相关系数与 $VR(q)$ 的联系。例如，考虑如下 AR(1) 形式的备择假设，

$$r_t = \phi r_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim i.i.d. (0, \sigma^2), \quad |\phi| < 1,$$

容易验证：⁹

$$\begin{aligned} VR(q) &= 1 + 2\sum_{k=1}^{q-1}\left(1 - \frac{k}{q}\right)\phi^k = 1 + 2\sum_{k=1}^{q-1}\phi^k - \frac{2}{q}\sum_{k=1}^{q-1}k\phi^k \\ &= 1 + 2\frac{\phi - \phi^q}{1 - \phi} - \frac{2}{q}\left[\frac{\phi - \phi^q}{(1 - \phi)^2} - (q-1)\frac{\phi^q}{1 - \phi}\right] \\ &= 1 + 2\frac{\phi}{1 - \phi} - \frac{2\phi(1 - \phi^q)}{q(1 - \phi)^2}. \end{aligned}$$

从上式可看出，当且仅当 $\phi = 0$ 时， $VR(q) = 1$ ；当 $\phi > 0$ 时， $VR(q) > 1$ 。

首先，本小节讨论如何利用 $VR(q)$ 的性质构造检验 RW1 的检验统计量。考虑如下原假设：

$$\mathbb{H}_0 : r_t = \mu + \varepsilon_t, \quad \varepsilon_t \sim i.i.d. \mathcal{N}(0, \sigma^2).$$

这里的正态假设不是必需的，可以放松为 ε_t 独立同分布且具有有限的四阶矩。考虑 $2T+1$ 个对数价格的观测值构成的样本 $\{p_0, p_1, \dots, p_{2T}\}$ 。利用此样本可以按照如下方式估计 μ 和 σ^2 ：

$$\hat{\mu} = \frac{1}{2T} \sum_{k=1}^{2T} (p_k - p_{k-1}) = \frac{1}{2T} (p_{2T} - p_0),$$

⁸ 在 RW5 条件下， $\text{Var}(r_t)$ 和 $\text{Var}(r_{t-1})$ 可能不相等。

⁹ 请参考本书第 2.3 节内容。

$$\hat{\sigma}^2 = \frac{1}{2T} \sum_{k=1}^{2T} (p_k - p_{k-1} - \hat{\mu})^2.$$

根据 Hayashi (2000) 的命题 1.5 可知, $\hat{\mu}$ 和 $\hat{\sigma}^2$ 分别是 μ 和 σ^2 的极大似然估计量, 而且, 由 Hayashi (2000) 的式 (1.5.16),

$$\sqrt{2T}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{d} \mathcal{N}(0, 2\sigma^4).$$

注意到, 在 RW1 条件下, $\text{Var}(r_t[2]) = 2\text{Var}(r_t)$, 于是, σ^2 还可以由下式估计:

$$\hat{\sigma}^2(2) = \frac{1}{2T} \sum_{k=1}^T (p_{2k} - p_{2k-2} - 2\hat{\mu})^2.$$

同样根据 Hayashi (2000) 的式 (1.5.16),

$$\sqrt{2T}(\hat{\sigma}^2(2) - \sigma^2) \xrightarrow{d} \mathcal{N}(0, 4\sigma^4).$$

在 \mathbb{H}_0 成立的情况下, Campbell et al. (1997) 证明:

$$\sqrt{2T}(\widehat{VR}(2) - 1) \xrightarrow{d} \mathcal{N}(0, 2),$$

其中, $\widehat{VR}(2) = \hat{\sigma}^2(2)/\hat{\sigma}^2$.

上述方差比检验可以容易地推广到多期对数收益率的情形。考虑 $qT + 1$ 个对数价格的观测值构成的样本 $\{p_0, p_1, \dots, p_{qT}\}$, 其中, q 为大于 1 的正整数, 以及基于此样本数据的如下估计量:

$$\begin{aligned}\hat{\mu} &= \frac{1}{Tq} \sum_{k=1}^{Tq} (p_k - p_{k-1}) = \frac{1}{Tq} (p_{Tq} - p_0), \\ \hat{\sigma}^2 &= \frac{1}{Tq} \sum_{k=1}^{Tq} (p_k - p_{k-1} - \hat{\mu})^2, \\ \hat{\sigma}^2(q) &= \frac{1}{Tq} \sum_{k=1}^T (p_{qk} - p_{qk-q} - q\hat{\mu})^2.\end{aligned}$$

可以证明 (参阅 Lo et al. (1988)):

$$\sqrt{qT}(\widehat{VR}(q) - 1) \xrightarrow{d} \mathcal{N}(0, 2(q-1)),$$

其中, $\widehat{VR}(q) = \hat{\sigma}^2(q)/\hat{\sigma}^2$.

为了改善检验的小样本性质, 常需要对上述检验统计量进行两个方向的重要修正: 一是在使用 q 期收益率估计方差时可以使用重叠的 q 期收益率, 即, 使用

$$\hat{\sigma}_{ol}^2(q) = \frac{1}{Tq^2} \sum_{k=q}^{Tq} (p_k - p_{k-q} - q\hat{\mu})^2$$

估计方差; 二是对方差估计量 $\hat{\sigma}^2$ 和 $\hat{\sigma}_{ol}^2(q)$ 的偏误进行修正, 具体地, 构造如下无偏方差估计量:

$$\begin{aligned}\bar{\sigma}^2 &= \frac{1}{Tq-1} \sum_{k=1}^{Tq} (p_k - p_{k-1} - \hat{\mu})^2, \\ \bar{\sigma}_{ol}^2(q) &= \frac{1}{m} \sum_{k=q}^{Tq} (p_k - p_{k-q} - q\hat{\mu})^2,\end{aligned}$$

其中,

$$m = q(Tq - q + 1) \left(1 - \frac{q}{Tq}\right).$$

Lo et al. (1988) 证明, 在 \mathbb{H}_0 成立的情况下,

$$\psi(q) \equiv \sqrt{Tq} (\overline{VR}(q) - 1) \left(\frac{2(2q-1)(q-1)}{3q} \right)^{-1/2} \xrightarrow{d} \mathcal{N}(0, 1),$$

其中, $\overline{VR}(q) = \bar{\sigma}_{ol}^2(q)/\bar{\sigma}^2$ 。

RW1 要求 ε_t 是独立同分布的, 但是资产收益率的波动可能具有时变性, 如果因为异方差的存在导致拒绝 RW1, 并不能说明 EMH 不成立。为此, 本小节接下来给出 $\overline{VR}(q)$ 在 RW5 条件下的抽样分布 (sampling distribution), 详细的证明过程可以参考 Lo et al. (1988)。

令 $r_t = \mu + \varepsilon_t$, 其中, ε_t 满足 Campbell et al. (1997) 中第 54 页的假设 H1-H4。这些假设在要求 r_t 不存在自相关的同时允许一般形式的异方差存在。记 δ_k 为 $\hat{\rho}_k$ 的渐近方差。若 H1-H4 成立, 则有:

$$\psi^*(q) \equiv \frac{\sqrt{Tq} (\overline{VR}(q) - 1)}{\sqrt{\hat{\theta}(q)}} \xrightarrow{d} \mathcal{N}(0, 1),$$

其中,

$$\begin{aligned} \hat{\theta}(q) &\equiv 4 \sum_{k=1}^{q-1} \left(1 - \frac{k}{q}\right)^2 \hat{\delta}_k, \\ \hat{\delta}_k &= \frac{Tq \sum_{j=k+1}^{Tq} (p_j - p_{j-1} - \hat{\mu})^2 (p_{j-k} - p_{j-k-1} - \hat{\mu})^2}{\left[\sum_{j=1}^{Tq} (p_j - p_{j-1} - \hat{\mu})^2 \right]^2}. \end{aligned}$$

随机游走检验: RW2

对时间跨度较长的时间序列数据来说, 假设不同时刻的收益率具有相同的分布是不切实际的。然而, 在不假设同分布的情况下去检验独立性并非易事。设想如果每一时刻的收益率的边缘分布都是不同的, 样本均值的抽样分布都是难以推导的。

实证研究中的两个方向可以看作是对 RW2 进行的一种“经济”意义上的检验, 它们分别是技术分析和过滤策略。技术分析建立在“历史会不断重演”的信仰之上, 认为价格、成交量及市场上的其他统计指标具有一定的规律性, 而且此类规律性往往可以总结为图形特征, 例如, 双底形态和头肩形态等。技术分析师通过研究这些形态特征推断未来的价格走势并从中获利。在 RW2 成立的条件下, 价格和成交量等历史信息对未来价格走势没有预测作用。EMH 意味着技术分析没有作用。如果某些图形形态或技术指标对未来的价格走势有预测作用, 那么就可以知道 RW2 是不成立的。由于技术分析较少使用严格的统计推断, 在本书中不再进行详细地介绍, 感兴趣的读者可以参考 Edwards et al. (2018); 陈学彬 (2021)。

为了检验 RW2, Alexander (1961) 提出了如下 $x\%$ 过滤策略: 如果某只股票日收盘价上涨至少 $x\%$, 买入并持有此股票, 直到它的收盘价从接下来的高点¹⁰下跌至少 $x\%$ 时卖

¹⁰ 高点是指上一次买入股票后遇到的最高收盘价。

出持有的股票并做空该股票；维持空头头寸，直到它的收盘价从接下来的低点¹¹上涨至少 $x\%$ 时填平空仓并买入此股票；循环上述操作，直到投资期结束，涨跌幅在 $x\%$ 以内的价格变动不触发交易。RW2 意味着不同时刻的对数收益率（即对数价格的变化）相互独立，若 RW2 成立， $x\%$ 过滤策略的实际收益将不会一直胜过简单的买入-持有策略。Fama et al. (1966) 的实证分析表明：在不考虑交易成本的情况下，非常小的滤子 x ($0.5\% - 1.5\%$) 确实可以带来超常收益；但是小的滤子会频繁触发交易，在考虑交易成本的情况下，即使往返交易成本小到 0.1% ，过滤策略的超常收益都将不复存在。

1.5 总结

假设对数收益率服从正态分布，或等价地，假设简单收益率服从对数正态分布，是许多金融理论研究的出发点，例如，Markowitz 的资产组合选择理论 (Markowitz (1952))、Sharpe 的 CAPM (Sharpe (1964)) 以及 Black & Scholes 的期权定价理论 (Scholes et al. (1973)) 等。然而，大量实证证据表明收益率具有厚尾和非对称等典型的非正态特征。本章给出了几种常用的刻画收益率非正态特征的概率分布，包括：稳定分布、(有偏) 学生 t 分布和混合正态分布。在本书第 14 章中，我们还将介绍广义极值分布 (GEV) 和广义帕累托分布 (GPD)，Malevergne et al. (2006) 详细地介绍了这两种分布，此外该书的第 2.2 节对收益率分布的历史做了较详尽的梳理。

本章的另外一个重要内容是有效市场假说和与其紧密联系的收益率随机游走的检验，写作过程中主要参考了经典的金融计量教材 Campbell et al. (1997)。关于随机游走检验的更详细的讨论，请参考 Linton (2019) 的第三章和第四章。这里需要说明的是随机游走检验和 EMH 检验并不是等价的。首先，随机游走检验只利用了资产价格的历史信息构造检验统计量。如果统计检验不能拒绝随机游走假设，虽然这对于 EMH 成立是有利的结果，但是这实际上仅仅表明：基于已有的价格数据，研究者未发现过去的价格变动能够预测未来的价格变动的证据，不能排除存在其他变量对未来的价格变动有预测作用的可能性；更重要的是，在前面讨论 RW 检验时，我们假设 $r_t = \mu + \varepsilon_t$ ，这里的常数 μ 为“正常”收益。一般地，对数收益率可以表示为 $r_t = \mu_t + \varepsilon_t$ ，其中， $\mu_t = \mathbb{E}[r_t | \mathcal{F}_{t-1}]$ 为期望收益率。所有的 EMH 检验都需要设定均衡模型进而定义期望收益率 μ_t 。因此，拒绝随机游走假设的原因可能确实是由于市场无效，但是也可能是因为均衡模型存在误设。“联合假设问题”的存在 (Fama (1970)) 使得 EMH 检验依然是一个悬而未决的话题。

在大数据时代，投资者可用来预测资产超额收益率的变量出现激增。与此同时，资产定价领域的学术研究发现的显著预测收益率的因子数量也越来越多。Cochrane (2011) 在摘要中写道：“Now we have a zoo of new factors”。如果金融市场的信息环境使得预测问题是高维的，那么金融市场理论模型中的投资者也应该面对着同样的高维问题。Martin et al. (2022) 将投资者建模为利用机器学习方法从历史数据中学习股票未来现金流与大量公司特征之间函数关系的贝叶斯主义者，研究发现：在均衡状态下，股票的定价满足收益率在样本外不可预测；然而，事后统计分析表明收益率在样本内具有很强的可预测性。

¹¹ 低点是指上一次卖出股票后遇到的最低收盘价。

附录 1: 对数正态分布的期望和方差

令 $y = \exp(x)$, 其中, $x \sim \mathcal{N}(\mu, \sigma^2)$ 。由对数正态分布的定义, y 服从对数正态分布。于是, y 期望为:

$$\begin{aligned}\mathbb{E}(y) &= \mathbb{E}[\exp(x)] = \int_{-\infty}^{\infty} e^x \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\ &= e^\mu \int_{-\infty}^{\infty} e^z \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2\sigma^2}\right) dz, \quad z = x - \mu \\ &= e^\mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2 - 2z\sigma^2}{2\sigma^2}\right) dz \\ &= e^{\mu+\sigma^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z-\sigma^2)^2}{2\sigma^2}\right) dz \\ &= e^{\mu+\sigma^2/2}.\end{aligned}$$

y 的方差为:

$$\begin{aligned}\text{Var}(y) &= \mathbb{E}(y^2) - [\mathbb{E}(y)]^2 \\ &= \mathbb{E}(y^2) - \exp(2\mu + \sigma^2),\end{aligned}$$

其中,

$$\begin{aligned}\mathbb{E}(y^2) &= \mathbb{E}(\exp(2x)) = \int_{-\infty}^{\infty} e^{2x} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\ &= \frac{1}{2} e^{2\mu} \int_{-\infty}^{\infty} e^{2(x-\mu)} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(2(x-\mu))^2}{2\sigma^2 \cdot 4}\right] d2(x-\mu) \\ &= e^{2\mu} \int_{-\infty}^{\infty} e^z \frac{1}{\sqrt{2\pi}\sigma^*} \exp\left[-\frac{z^2}{2\sigma^{*2}}\right] dz, \quad z = 2(x-\mu), \sigma^* = 2\sigma \\ &= e^{2\mu+\sigma^{*2}/2} = e^{2\mu+2\sigma^2}.\end{aligned}$$

因此, 我们有:

$$\text{Var}(y) = \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2).$$

注 我们也可以利用正态分布函数的矩生成函数推导上述结论, 关于矩生成函数的介绍请参考 [Casella et al. \(2002\)](#) 的第 2.3 节。

附录 2: 正态性检验

金融时间序列数据通常具有尖峰、厚尾和非对称等典型的非正态特征 (见 [Jondeau et al. \(2007\)](#))。我们这里介绍几种检验正态性的统计方法及其 R 语言实现。

可视化方法

常用的识别正态性是否满足的可视化方法有直方图 (histogram) 和 Q-Q 图。

假设随机变量 x 的密度函数是 $f(x)$, f 的支撑¹² (support) 是 $[a, b]$ 。考虑 $[a, b]$ 的一

¹² 随机变量 x 的概率密度函数 f 的支撑为 $\{x \in \mathbb{R} : f(x) \neq 0\}$ 。

个划分 (partition) $a = t_0 < t_1 < \dots < t_m = b$, 直方图通过如下方式估计 x 的概率密度:

$$\hat{f}(x) = \frac{n_i}{n(t_{i+1} - t_i)}, \quad t_i \leq x < t_{i+1},$$

其中, n 为样本量, n_i 是落入区间 $[t_i, t_{i+1})$ 的样本个数。

若随机变量 x 服从 $\mathcal{N}(\mu, \sigma^2)$ 分布, 则 x 的 α 分位数 $Q(\alpha) = \mu + \sigma z_\alpha$, 其中, z_α 是标准正态分布的 α 分位数。因此, 以 α 作为横轴绘制 $Q(\alpha)$ 的图像将得到一条截距为 μ 斜率为 σ 的直线。给定样本 $\{x_1, x_2, \dots, x_n\}$, 我们可以将样本分位数和同阶的正态分位数分别作为纵坐标和横坐标绘制散点图, 若数据服从正态分布, 散点图应该大致是一条直线。这便是 Q-Q 图的基本思想。

我们以上证指数从 2007 年 4 月 13 日至 2019 年 12 月 31 日的周对数收益率为例说明如何利用 R 语言绘制直方图和 Q-Q 图, 使用的核心代码如下:

```

1 #读入上证指数数据, 计算周对数收益率rtn
2 #完整代码可从本书网站下载, 这里只列出画图的核心代码
3 mu <- mean(rtn)
4 sd <- sd(rtn)
5 #从rtn的最小值到最大值, 生成间隔0.001的数列
6 nrx <- seq(min(rtn), max(rtn), by=0.001)
7 #计算数列nrx对应的正态pdf取值,
8 #正态分布期望是mu, 标准差是sd, 由数据rtn计算得出
9 nry <- dnorm(nrx, mu, sd)
10 #将nrx和nry按列合并, 并转换为data frame
11 normdata <- as.data.frame(cbind(nrx, nry))
12 library(tidyverse)
13 #直方图
14 rtn %>% # pipe operator
15   ggplot(aes(rtn)) +
16   geom_histogram(aes(y = ..density..), fill="yellow", colour="grey30") +
17   labs(x = "weekly returns") +
18   #在直方图上加正态pdf曲线
19   geom_line(mapping = aes(x=nrx, y=nry), data = normdata, colour="blue")
20 #Q-Q图
21 ggplot(rtn, aes(sample=rtn)) +
22   labs(x = "Normal quantile", y="Quantile returns")+
23   stat_qq(colour="blue") +
24   stat_qq_line(colour="red")

```

上述程序的运行结果如图 1.15 和 1.16 所示, 从图中我们可以看出上证指数收益率具有尖峰和厚尾的特征, Q-Q 图显示其分布在左右尾都厚于正态分布。

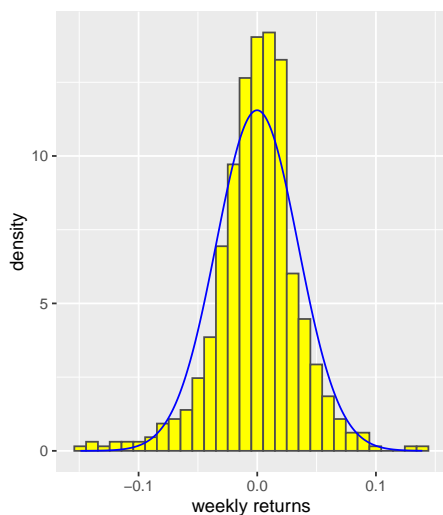


图 1.15: 上证指数收益率直方图

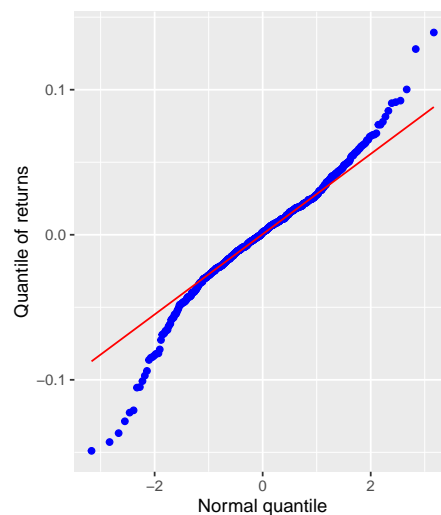


图 1.16: 上证指数收益率 Q-Q 图

我们也可以使用 R 的 `qqtest` 程序包提供的 `qqtest` 命令绘制有置信区间的 Q-Q 图：

```
1 library(qqtest)
2 qqtest(rtn,main="Returns")
```

运行结果如图 1.17 所示。

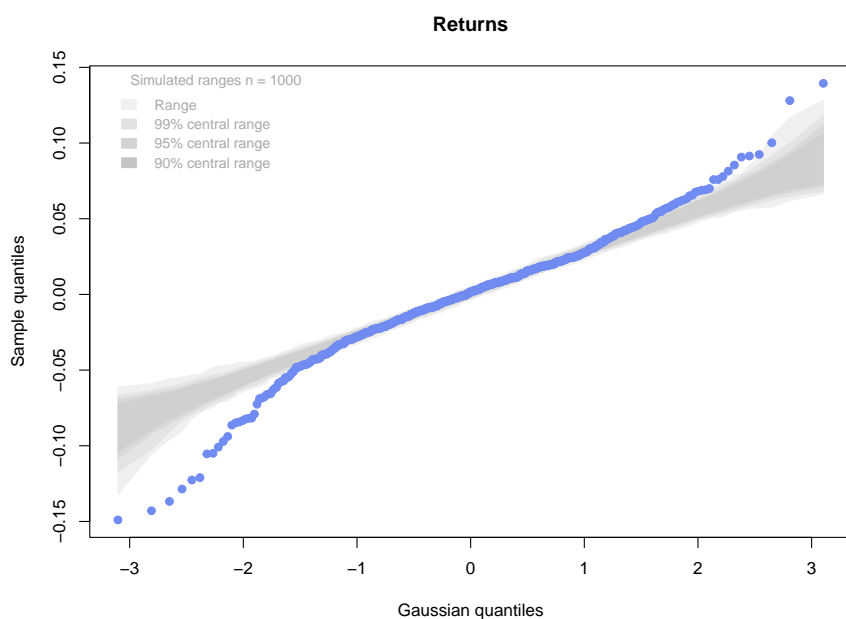


图 1.17: 上证指数收益率 Q-Q 图

峰度和偏度

我们还可以利用偏度和峰度判断数据是否服从正态分布。偏度的定义为：

$$s = \mathbb{E} \left[\left(\frac{x - \mu}{\sigma} \right)^3 \right];$$

峰度的定义为：

$$\kappa = \mathbb{E} \left[\left(\frac{x - \mu}{\sigma} \right)^4 \right].$$

偏度用来衡量随机变量分布的非对称性。若偏度为负，说明发生绝对值较大的负的实现值的可能性要大于发生绝对值较大的正的实现值的可能性。峰度用来衡量随机变量的厚尾程度。可以证明正态分布的峰度为 3，若随机变量的峰度大于 3，说明绝对值较大的实现值发生的可能性高于正态分布的预期。

给定样本 $\{x_1, x_2, \dots, x_n\}$ ，我们可以计算样本偏度和样本峰度，它们的定义分别为：

$$\hat{s} = \frac{1}{(n-1)\hat{\sigma}^3} \sum_{i=1}^n (x_i - \hat{\mu})^3$$

和

$$\hat{\kappa} = \frac{1}{(n-1)\hat{\sigma}^4} \sum_{i=1}^n (x_i - \hat{\mu})^4,$$

其中， $\hat{\mu}$ 为样本均值， $\hat{\sigma}$ 为样本标准差。在正态分布假设下，可以证明：

$$\sqrt{n}\hat{s} \xrightarrow{d} \mathcal{N}(0, 6), \quad \sqrt{n}(\hat{\kappa} - 3) \xrightarrow{d} \mathcal{N}(0, 24).$$

因为正态分布的偏度为 0，为了检验正态性，我们可以考虑检验 $\mathbb{H}_0: s = 0$ ，其备择假设为 $\mathbb{H}_1: s \neq 0$ 。基于样本偏度的检验统计量为：

$$t = \frac{\hat{s}}{\sqrt{6/n}},$$

此统计量渐近服从标准正态分布。类似地，我们也可以通过原假设为 $\mathbb{H}_0: k - 3 = 0$ 的双边检验判断数据是否服从正态分布，其检验统计量为：

$$t = \frac{\hat{\kappa} - 3}{\sqrt{24/n}},$$

此统计量也渐近服从标准正态分布。Jarque & Bera 的正态检验（见 Jarque et al. (1980)）将上述两个检验统计量结合，其检验统计量的形式为：

$$JB = \frac{\hat{s}^2}{6/n} + \frac{(\hat{\kappa} - 3)^2}{24/n},$$

可以证明 $JB \xrightarrow{d} \chi^2(2)$ 。

在 R 中，我们可以用如下代码对上证指数收益率进行 JB 正态检验：

```
1 #JB test
2 library(fBasics)
3 normalTest(rtn,method='jb')
4 #下面是检验的输出结果
5 Title:
6 Jarque - Bera Normalality Test
7 Test Results:
8 STATISTIC:
9 X-squared: 181.7368
10 P VALUE:
```

11 Asymptotic p Value: < 2.2e-16

从输出结果看，JB 检验的 p 值近似为 0，我们可以拒绝上证指数收益率服从正态分布的原假设。

统计检验

除了 JB 检验之外，常用的正态性检验还有 Shapiro-Wilk 检验、Kolmogorov-Smirnov 检验、Cramér-von Mises 检验和 Anderson-Darling 检验（见 Serfling (1980)）。

Shapiro-Wilk 检验（见 Shapiro et al. (1965)）统计量定义为：

$$SW = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

其中， $x_{(i)}$ 是第 i 个样本次序统计量， \bar{x} 是样本均值。权重 a_i 的计算方式如下：令 $\{y_1, \dots, y_n\}$ 是来自标准正态分布的一组相互独立的随机变量， $y_{(i)}$ 是第 i 个次序统计量。定义次序统计量的期望向量为 $\mathbf{m}' = (m_1, \dots, m_n)$ ，其中， $m_i = \mathbb{E}(y_{(i)})$ 。再定义次序统计量的协方差矩阵为 \mathbf{V} ，其 (i, j) 位置上的元素为 $v_{ij} = \text{Cov}(y_{(i)}, y_{(j)})$ 。权重向量 $\mathbf{a}' = \{a_1, \dots, a_n\} = \mathbf{m}'\mathbf{V}^{-1}[(\mathbf{m}'\mathbf{V}^{-1})(\mathbf{V}^{-1}\mathbf{m})]^{-1/2}$ 。

Kolmogorov-Smirnov 检验统计量定义为：

$$KS = \sqrt{n} \sup_x |F_n(x) - F_0(x)|,$$

其中， $F_n(x)$ 是经验分布函数；若原假设为数据服从 $\mathcal{N}(\mu, \sigma^2)$ 分布，则 $F_0(x) = \Phi(\frac{x - \mu}{\sigma})$ ， $\Phi(\cdot)$ 为标准正态的累积分布函数。

Cramér-von Mises 检验统计量定义为：

$$CM = n \cdot \int_{-\infty}^{\infty} (F_n(x) - F_0(x))^2 dF_0(x).$$

Anderson-Darling 检验（见 Anderson et al. (1954)）统计量定义为：

$$AD = n \cdot \int_{-\infty}^{\infty} \frac{(F_n(x) - F_0(x))^2}{F_0(x)(1 - F_0(x))} dF_0(x).$$

由上式我们可以看出，Anderson-Darling 检验统计量对处在分布尾部的观测值给予了更大的权重。关于正态性检验更全面的介绍以及各个检验统计量的小样本表现读者可以参考 Thadewald et al. (2007)。

在 R 中，我们可以通过下列代码对上证指数收益率进行上述四种检验：

```
1 #Shapiro-Wilk test
2 shapiro.test(rtn)
3 #Anderson-Darling test
4 nortest::ad.test(rtn)
5 #Cramer-von Mises test
6 nortest::cvm.test(rtn)
7 #Kolmogorov-Smirnov test
8 nortest::lillie.test(rtn)
```

四个检验的 p 值都近似为 0，均拒绝上证指数收益率服从正态的原假设。

多元正态检验

因为构成多元正态分布的每一个随机变量都服从一元正态分布，所以我们可以通过一元正态的检验判断随机变量向量是否服从多元正态分布。但是，多个一元正态检验的真实显著性水平与多元正态检验的名义显著性水平不相等，计算检验的 p 值时需要做调整，可以采用 R 中的 `stats::p.adjust` 命令。

多元正态分布的一个重要性质是构成多元正态分布的随机变量的任意线性组合都服从一元正态分布。我们可以随机生成 m 个单位向量，进而构造数据的 m 个线性组合，通过检验这 m 个线性组合是否服从一元正态分布来判断数据是否服从多元正态分布。本书网站提供了用此方法检验二元正态分布的 R 函数 `testnormality.R`，函数返回的是检验的 p 值。

我们这里再介绍两种检验多元正态分布的可视化方法。因为服从多元正态分布的多个随机变量如果不相关则一定相互独立，所以若发现两个随机变量之间有非线性的关系，则它们一定不服从多元正态分布。因此，我们可以通过绘制变量两两之间的散点图观测不同变量的数据之间是否具有非线性关系，从而判断数据是否服从多元正态分布。在 R 中，可以使用 `GGally` 程序包的 `ggpairs` 命令绘制散点图。

此外，给定 p 维随机向量 \mathbf{y} 的样本 $\{\mathbf{y}_1, \dots, \mathbf{y}_n\}$ ，若 \mathbf{y} 服从多元正态分布，平方马氏距离 $d_i^2 = (\mathbf{y}_i - \bar{\mathbf{y}})' \mathbf{S}^{-1} (\mathbf{y}_i - \bar{\mathbf{y}})$ 近似服从 $\chi^2(p)$ 分布（见 Johnson et al. (2007) 的第 184 页），其中， $\bar{\mathbf{y}}$ 是样本均值， \mathbf{S} 是样本协方差矩阵。因此，我们可以通过 $\{d_1^2, \dots, d_n^2\}$ 和卡方分布的 Q-Q 图判断数据是否服从多元正态分布。下面我们来自二元正态分布的模拟数据说明如何在 R 中实现此可视化方法：

```
1 library(MASS)
2 mu      = c(0,1) #二元正态的期望向量
3 Sig     = matrix(c(1, 0.6, 0.6, 1), 2, 2) #二元正态的协方差矩阵
4 n       = 5000 #样本量
5 set.seed(1234)
6 biv     = MASS::mvrnorm(n, mu, Sig) #生成二元正态随机数
7 Sbiv    = cov(biv) #样本协方差矩阵
8 D2      = mahalanobis(biv, colMeans(biv), Sbiv) #计算马氏距离的平方
9 D2 <- as.data.frame(D2)
10 ggplot(D2, aes(sample=D2)) +
11   labs(x = "Theoretical quantile", y="Quantile of Distance")+
12   #dparams是卡方分布的自由度
13   stat_qq(distribution = stats::qchisq, dparams=2, colour="blue") +
14   stat_qq_line(distribution = stats::qchisq, dparams=2, colour="red")
```

运行结果如图 1.18 所示，Q-Q 图近似为一条直线。

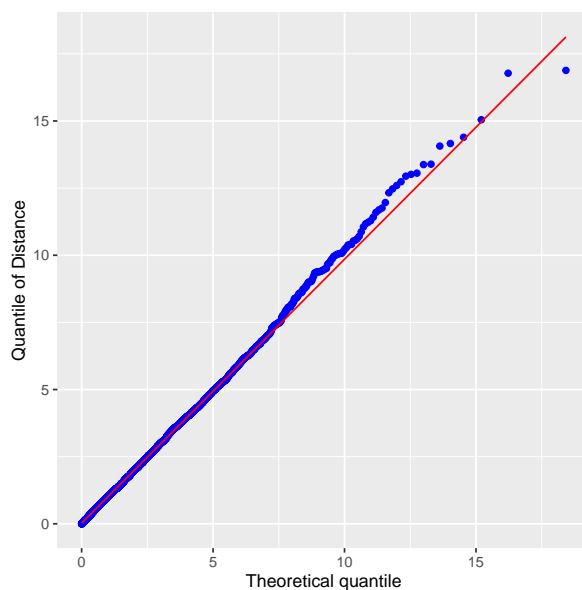


图 1.18: 二元正态分布平方马氏距离的 Q-Q 图

第一章习题

1. 实证：下载任意一支中国股票的价格数据，计算其日度、周度和月度的简单收益率和对数收益率，并通过统计检验或可视化的方法分析该股票收益率数据的分布特征。
2. 实证：下载任意一支中国股票的价格数据，计算其日度、周度和月度的对数收益率，并分析是否随着收益率频率的降低其分布更加接近于正态分布。
3. 实证：下载上证指数的日收盘价数据，对其进行同向/反向比检验和方差比检验。
4. 实证：下载上证指数的日收盘价数据，对其进行过滤策略检验，分析滤子大小及交易成本对检验结果的影响。（参考 Fama et al. (1966) 或 姜近勇，潘冠中 (2011) 中 5.3.2 小节给出的具体步骤）

第二部分 线性时间序列模型

第二章 一元线性时间序列模型

Hamilton (1994); Martin et al. (2013)

2.1 自相关函数

2.2 移动平均模型

2.3 自回归模型

2.4 自回归移动平均模型

第三章 多元线性时间序列模型

Tsay (2013)

第四章 单位根与协整

单位根检验和随机游走检验的区别

第五章 状态空间模型及卡尔曼滤波*

第六章 混频数据抽样回归

Mixed Frequency Data Sampling Regression MIDAS

第三部分 非线性模型

第七章

第八章

第九章 深度学习模型

第四部分 波动率模型

第十章 一元波动率模型

第十一章 多元波动率模型

第十二章 随机波动率模型*

第五部分 量化风险管理

第十三章 单资产风险测度

第十四章 极值理论*

第十五章 Copula

McNeil et al. (2015)

Chabi-Yo et al. (2018)

第十六章 系统风险测度*

参考文献

- 陈公宁, 2007. 矩阵理论与应用 [M]. 北京: 科学出版社.
- 陈学彬, 2021. Python 量化投资基础教程 [M]. 北京: 高等教育出版社.
- 方颖, 赵扬, 2011. 寻找制度的工具变量: 估计产权保护对中国经济增长的贡献 [J]. 经济研究(5): 138-148.
- 姜近勇, 潘冠中, 2011. 金融计量学 [M]. 北京: 中国财政经济出版社.
- 李高荣, 吴密霞, 2021. 多元统计分析 [M]. 北京: 科学出版社.
- 李航, 2022. 机器学习方法 [M]. 北京: 清华大学出版社.
- 李勉群 (美), 苏子英 (美) 【著】, 张然 【译】, 2019. 阿尔法经济学: 赢取资本超额收益的法则 [M]. 北京: 北京大学出版社.
- 刘浩洋, 户将, 李勇锋, 文再文, 2020. 最优化: 建模、算法与理论 [M]. 北京: 高等教育出版社.
- 苏淳, 冯群强, 2020. 概率论 (第三版) [M]. 北京: 科学出版社.
- 王松桂, 史建红, 尹素菊, 吴密霞, 2004. 线性模型引论 [M]. 北京: 科学出版社.
- ABADIE A, ATHEY S, IMBENS G W, et al., 2020. Sampling-based versus design-based uncertainty in regression analysis[J]. *Econometrica*, 88(1): 265-296.
- ABADIE A, ATHEY S, IMBENS G W, et al., 2022. When should you adjust standard errors for clustering?[J]. *arXiv preprint arXiv:1710.02926v3*.
- ABADIR K M, MAGNUS J R, 2005. Matrix algebra[M]. New York: Cambridge University Press.
- ADRIAN T, BRUNNERMEIER M K, 2016. CoVaR[J]. *American Economic Review*, 106(7): 1705-1741.
- ALEXANDER S S, 1961. Price movements in speculative markets: Trends or random walks[J]. *Industrial Management Review*, 2(2): 7-26.
- AMEMIYA T, 1977. A note on a heteroscedastic model[J]. *Journal of Econometrics*, 6(3): 365-370.
- ANDERSON T W, DARLING D A, 1954. A test of goodness of fit[J]. *Journal of the American Statistical Association*, 49(268): 765-769.
- ANDREWS D W, 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation[J]. *Econometrica*: 817-858.
- ANGRIST J D, PISCHKE J S, 2008. Mostly harmless econometrics: An empiricist's companion[M]. Princeton: Princeton University Press.
- ANSELIN L, 1988. Spatial econometrics: methods and models[M]. Netherlands: Kluwer Academic Publishers.
- BACHELIER L, 1900. Theory of speculation[M]//COOTNER P H. The random character of stock market prices. Cambridge: MIT Press, 1964.
- BAI J, 2009. Panel data models with interactive fixed effects[J]. *Econometrica*, 77(4): 1229-1279.
- BALI T G, ENGLE R F, MURRAY S, 2016. Empirical asset pricing: The cross section of stock returns[M]. New Jersey: John Wiley & Sons.
- BELL R M, MCCAFFREY D F, 2002. Bias reduction in standard errors for linear regression with multi-stage samples[J]. *Survey Methodology*, 28(2): 169-182.
- BELSLEY D A, KUH E, WELSCH R E, 2004. Regression diagnostics: Identifying influential data and sources of collinearity[M]. New Jersey: John Wiley & Sons.
- BERKOWITZ D, CANER M, FANG Y, 2012. The validity of instruments revisited[J]. *Journal of Econometrics*, 166(2): 255-266.
- BISHOP C M, 2006. Pattern recognition and machine learning[M]. New York: Springer.
- BOX G E, PIERCE D A, 1970. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models[J]. *Journal of the American Statistical Association*, 65(332): 1509-1526.
- BREDON G E, 2013. Topology and geometry[M]. New York: Springer.
- BREUSCH T S, 1978. Testing for autocorrelation in dynamic linear models[J]. *Australian economic papers*, 17(31): 334-355.

- BREUSCH T S, PAGAN A R, 1979. A simple test for heteroscedasticity and random coefficient variation[J]. *Econometrica*: 1287-1294.
- BROCKWELL P J, DAVIS R A, 1991. Time series: theory and methods[M]. 2nd ed. New York: Springer Science & Business Media.
- CAI Z, FAN J, YAO Q, 2000. Functional-coefficient regression models for nonlinear time series[J]. *Journal of the American Statistical Association*, 95(451): 941-956.
- CALAFIORE G C, EL GHAOU L, 2014. Optimization models[M]. Cambridge: Cambridge university press.
- CAMERON A C, MILLER D L, 2015. A practitioner's guide to cluster-robust inference[J]. *Journal of Human Resources*, 50(2): 317-372.
- CAMERON A C, TRIVEDI P K, 2005. Microeconometrics: methods and applications[M]. New York: Cambridge University Press.
- CAMERON A C, GELBACH J B, MILLER D L, 2008. Bootstrap-based improvements for inference with clustered errors[J]. *Review of Economics and Statistics*, 90(3): 414-427.
- CAMPBELL J Y, LO A W, MACKINLAY A, 1997. The econometrics of financial markets[M]. Princeton: Princeton University Press.
- CASELLA G, BERGER R L, 2002. Statistical inference[M]. California: Thomson Learning.
- CERULLI G, 2015. Econometric evaluation of socio-economic programs theory and applications[M]. Berlin: Springer.
- CHABI-YO F, RUENZI S, WEIGERT F, 2018. Crash sensitivity and the cross section of expected stock returns[J]. *Journal of Financial and Quantitative Analysis*, 53(3): 1059-1100.
- COCHRANE J H, 2005. Asset pricing: Revised edition[M]. Princeton: Princeton University Press.
- COCHRANE J H, 2011. Presidential address: Discount rates[J]. *Journal of Finance*, 66(4): 1047-1108.
- COWLES A, JONES H E, 1937. Some a posteriori probabilities in stock market action[J]. *Econometrica*: 280-294.
- CUNNINGHAM S, 2021. Causal inference: The mixtape[M]. New Haven: Yale University Press.
- DAVIDSON R, FLACHAIRE E, 2008. The wild bootstrap, tamed at last[J]. *Journal of Econometrics*, 146(1): 162-169.
- DUFLO E, DUPAS P, KREMER M, 2011. Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya[J]. *American Economic Review*, 101(5): 1739-74.
- EDWARDS R D, MAGEE J, BASSETTI W C, 2018. Technical analysis of stock trends[M]. 11st ed. New York: Routledge.
- EFRON B, 1979. Bootstrap Methods: Another Look at the Jackknife[J]. *Annals of Statistics*, 7(1): 1 - 26.
- EFRON B, 1982. The Jackknife, the Bootstrap and other resampling plans[M]. Philadelphia: Society for Industrial and Applied Mathematics.
- EFRON B, 1987. Better Bootstrap confidence intervals[J]. *Journal of the American Statistical Association*, 82(397): 171-185.
- EFRON B, HASTIE T, 2016. Computer age statistical inference: Algorithms, evidence, and data science[M]. New York: Cambridge University Press.
- FAMA E F, 1963. Mandelbrot and the stable Paretian hypothesis[J]. *Journal of Business*, 36(4): 420-429.
- FAMA E F, 1970. Efficient capital markets: A review of theory and empirical work[J]. *Journal of Finance*, 25(2): 383-417.
- FAMA E F, BLUME M E, 1966. Filter rules and stock-market trading[J]. *Journal of Business*, 39(1): 226-241.
- FAMA E F, FRENCH K R, 1992. The cross-section of expected stock returns[J]. *Journal of Finance*, 47(2): 427-465.
- FAN J, YAO Q, 2015. The elements of financial econometrics[M]. Beijing: Science Press.
- FERNÁNDEZ C, STEEL M F, 1998. On Bayesian modeling of fat tails and skewness[J]. *Journal of the American Statistical Association*, 93(441): 359-371.
- GODFREY L G, 1978. Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables[J]. *Econometrica*: 1293-1301.

- GREENE W H, 2018. *Econometric analysis*[M]. 8th ed. New York: Pearson Education.
- GROSSMAN S J, STIGLITZ J E, 1980. On the impossibility of informationally efficient markets[J]. *American Economic Review*, 70(3): 393-408.
- HALL P, HEYDE C C, 1980. *Martingale limit theory and its application*[M]. New York: Academic press.
- HAMILTON J D, 1994. *Time series analysis*[M]. New Jersey: Princeton University Press.
- HANSEN B E, 1992. Consistent covariance matrix estimation for dependent heterogeneous processes[J]. *Econometrica*: 967-972.
- HANSEN B E, 2022a. *Econometrics*[M]. Princeton: Princeton University Press.
- HANSEN B E, 2022b. *Probability and statistics for economists*[M]. Princeton: Princeton University Press.
- HANSEN B E, 2022c. A modern Gauss-Markov theorem[J]. forthcoming in *Econometrica*.
- HASTIE T, ROBERT T, FRIEDMAN J, 2016. *The elements of statistical learning: Data mining, inference, and prediction*[M]. 2nd ed. Berlin: Springer.
- HAYASHI F, 2000. *Econometrics*[M]. Princeton: Princeton University Press.
- HINKLEY D V, 1977. Jackknifing in unbalanced situations[J]. *Technometrics*, 19(3): 285-292.
- HONG Y, 2020. *Foundations of modern econometrics: A unified approach*[M]. London: World Scientific.
- HORN R A, JOHNSON C R, 2013. *Matrix analysis*[M]. New York: Cambridge University Press.
- HORN S D, HORN R A, DUNCAN D B, 1975. Estimating heteroscedastic variances in linear models[J]. *Journal of the American Statistical Association*, 70(350): 380-385.
- HOROWITZ J L, 2019. Bootstrap methods in econometrics[J]. *Annual Review of Economics*, 11: 193-224.
- HUANG D, JIANG F, TU J, et al., 2015. Investor sentiment aligned: A powerful predictor of stock returns[J]. *Review of Financial Studies*, 28(3): 791-837.
- HURN S, MARTIN V, PHILLIPS P C, et al., 2021. *Financial econometric modeling*[M]. New York: Oxford University Press.
- INGERSOLL J E, 1987. *Theory of financial decision making*[M]. Maryland: Rowman & Littlefield.
- JARQUE C M, BERA A K, 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals[J]. *Economics Letters*, 6(3): 255-259.
- JOHNSON R A, WICHERN D W, 2007. *Applied multivariate statistical analysis*[M]. 6th ed. London: Pearson.
- JONDEAU E, POON S H, ROCKINGER M, 2007. *Financial modeling under non-Gaussian distributions*[M]. London: Springer Science & Business Media.
- KOENKER R, 1981. A note on studentizing a test for heteroscedasticity[J]. *Journal of Econometrics*, 17(1): 107-112.
- KOENKER R, 2005. *Quantile regression*[M]. New York: Cambridge University Press.
- LI Q, RACINE J S, 2007. *Nonparametric econometrics: theory and practice*[M]. New Jersey: Princeton University Press.
- LIN Z, BAI Z, 2010. *Probability inequalities*[M]. Beijing: Science Press Beijing and Springer.
- LINTON O, 2019. *Financial econometrics: Models and methods*[M]. Padstow: Cambridge University Press.
- LJUNG G M, BOX G E, 1978. On a measure of lack of fit in time series models[J]. *Biometrika*, 65(2): 297-303.
- LO A W, MACKINLAY A C, 1988. Stock market prices do not follow random walks: Evidence from a simple specification test[J]. *Review of Financial Studies*, 1(1): 41-66.
- MACKINNON J G, 2006. Bootstrap methods in econometrics[J]. *Economic Record*, 82: S2-S18.
- MACKINNON J G, WHITE H, 1985. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties[J]. *Journal of Econometrics*, 29(3): 305-325.
- MACKINNON J G, NIELSEN M Ø, WEBB M D, 2022a. Cluster-robust inference: A guide to empirical practice[J]. *Journal of Econometrics*.
- MACKINNON J G, NIELSEN M Ø, WEBB M D, 2022b. Fast and reliable jackknife and bootstrap methods for cluster-robust inference[R]. Queen's University: Queen's Economics Department Working Paper No.1485.
- MACKINNON J G, NIELSEN M Ø, WEBB M D, 2022c. Leverage, influence, and the jackknife in clustered regression models: Reliable inference using `summclust`[J]. arXiv preprint arXiv:2205.03288.

- MALEVERGNE Y, SORNETTE D, 2006. Extreme financial risks: From dependence to risk management[M]. Netherlands: Springer.
- MAMMEN E, 1993. Bootstrap and wild bootstrap for high dimensional linear models[J]. *Annals of Statistics*, 21(1): 255-285.
- MANDELBORT B, 1963. The variation of certain speculative prices[J]. *Journal of Business*, 36(4): 394-419.
- MARKOWITZ H, 1952. Portfolio selection[J]. *Journal of Finance*, 7(1): 77-91.
- MARRON J S, WAND M P, 1992. Exact mean integrated squared error[J]. *Annals of Statistics*, 20(2): 712-736.
- MARTIN I W, NAGEL S, 2022. Market efficiency in the age of big data[J]. *Journal of Financial Economics*, 145(1): 154-177.
- MARTIN V, HURN S, HARRIS D, 2013. Econometric modelling with time series: specification, estimation and testing[M]. New York: Cambridge University Press.
- MCNEIL A J, FREY R, EMBRECHTS P, 2015. Quantitative risk management: concepts, techniques and tools[M]. Rev. ed. New Jersey: Princeton University Press.
- MURPHY K P, 2022. Probabilistic machine learning: An introduction[M]. Cambridge: MIT Press.
- MURPHY K P, 2023. Probabilistic machine learning: Advanced topics[M]. Cambridge: MIT Press.
- NAGEL S, 2021. Machine learning in asset pricing[M]. New Jersey: Princeton University Press.
- NEWKEY W K, WEST K D, 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix[J]. *Econometrica*, 55(3): 703-708.
- PENROSE R, 1955. A generalized inverse for matrices[J]. *Mathematical Proceedings of the Cambridge Philosophical Society*, 51(3): 406-413.
- PESARAN M H, 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure[J]. *Econometrica*, 74(4): 967-1012.
- PETERSEN M A, 2009. Estimating standard errors in finance panel data sets: Comparing approaches[J]. *Review of Financial Studies*, 22(1): 435-480.
- PÖTSCHER B M, PREINERSTORFER D, 2022. A Modern Gauss-Markov Theorem? Really?[J]. arXiv preprint arXiv:2203.01425.
- PUSTEJOVSKY J E, TIPTON E, 2018. Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models[J]. *Journal of Business & Economic Statistics*, 36(4): 672-683.
- RACHEV S T, MENN C, FABOZZI F J, 2005. Fat-tailed and skewed asset return distributions: Implications for risk management, portfolio selection, and option pricing[M]. New Jersey: John Wiley & Sons.
- ROBERTS M R, WHITED T M, 2013. Endogeneity in empirical corporate finance[M]//CONSTANTINIDES G M, HARRIS M, STULZ R M. *Handbook of the Economics of Finance*: 2A. Amsterdam: Elsevier.
- SCHOLES M, BLACK F, 1973. The pricing of options and corporate liabilities[J]. *Journal of Political Economy*, 81(3): 637-654.
- SERFLING R J, 1980. Approximation theorems of mathematical statistics[M]. New York: John Wiley & Sons.
- SHAO J, TU D, 1995. The Jackknife and Bootstrap[M]. New York: Springer.
- SHAPIRO S S, WILK M B, 1965. An analysis of variance test for normality (complete samples)[J]. *Biometrika*, 52: 591-611.
- SHARPE W F, 1964. Capital asset prices: A theory of market equilibrium under conditions of risk[J]. *Journal of Finance*, 19(3): 425-442.
- SHLEIFER A, VISHNY R W, 1997. The limits of arbitrage[J]. *Journal of Finance*, 52(1): 35-55.
- STOCK J H, WATSON M W, 2020. Introduction to Econometrics[M]. 4 (Global Edition) ed. New York: Pearson.
- STRANG G, 2016. Introduction to linear algebra[M]. 5th ed. Massachusetts: Wellesley-Cambridge Press.
- THADEWALD T, BÜNING H, 2007. Jarque-Bera test and its competitors for testing normality—a power comparison[J]. *Journal of Applied Statistics*, 34(1): 87-105.
- TSAY R S, 2013. Multivariate time series analysis: with r and financial applications[M]. New Jersey: John Wiley & Sons.

- TSAY R S, 2014. An introduction to analysis of financial data with R[M]. New Jersey: John Wiley & Sons.
- VAN DER VAART A W, 2000. Asymptotic statistics[M]. Cambridge: Cambridge university press.
- VERBEEK M, 2021. Panel methods for finance: A guide to panel data econometrics for financial applications[M]. Berlin: De Gruyter.
- WHITE H, 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity[J]. *Econometrica*: 817-838.
- WHITE H, 2001. Asymptotic theory for econometricians[M]. Rev. ed. California: Academic Press.
- WHITE H, DOMOWITZ I, 1984. Nonlinear regression with dependent observations[J]. *Econometrica*: 143-161.
- WOOLDRIDGE J M, 2010. Econometric analysis of cross section and panel data[M]. Cambridge: The MIT Press.
- WOOLDRIDGE J M, 2019. Introductory econometrics: a modern approach[M]. 7th ed. Boston: Cengage Learning.
- WU C F J, 1986. Jackknife, bootstrap and other resampling methods in regression analysis[J]. *Annals of Statistics*, 14(4): 1261-1295.
- XU Q H, CAI Z W, FANG Y, 2016. Panel data models with cross-sectional dependence: A selective review[J]. *Applied Mathematics-A Journal of Chinese Universities*, 31(2): 127-147.

附录

附录 A EM 算法 *

附录 B 贝叶斯估计 *

内容提要



贝叶斯估计提供了一个将先验知识纳入统计估计的框架。